

GenCore version 5.1.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 7, 2004, 15:17:31 ; Search time 109 Seconds
(without alignments)
5865.169 Million cell updates/sec

Title: US-09-976-782-15
Perfect score: 1152
Sequence: 1 cgcattggagccgctagcc.....aagcagagatgaacacaggg 1152

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 682709 seqs, 277475446 residues

Total number of hits satisfying chosen parameters: 1365418

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents NA:*
1: /cgn2_6/ptodata/2/ina/5A_COMB.seq:*
2: /cgn2_6/ptodata/2/ina/5B_COMB.seq:*
3: /cgn2_6/ptodata/2/ina/6A_COMB.seq:*
4: /cgn2_6/ptodata/2/ina/6B_COMB.seq:*
5: /cgn2_6/ptodata/2/ina/PCTUS_COMB.seq:*
6: /cgn2_6/ptodata/2/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	729.8	63.4	2226	2	US-08-031-538-1
2	484.4	42.0	1073	1	US-08-356-405-8
3	424.6	36.9	1686	1	US-08-356-405-1
4	416	36.1	2155	2	US-08-031-538-3
5	104.4	9.1	1382	4	US-09-016-434-1256
6	95.2	8.3	1984	4	US-09-016-434-1077
7	94	8.2	1756	4	US-09-016-434-1197
8	93	8.1	2428	3	US-08-475-742-15
9	93	8.1	2428	4	US-08-261-293-15
10	93	8.1	2818	4	US-09-376-594-629
11	91.2	7.9	1567	1	US-08-722-001-24
12	89.8	7.8	1227	1	US-08-351-473B-7
13	88.8	7.7	1738	1	US-08-334-698-3
14	88.8	7.7	1738	1	US-08-228-932-3
15	88.8	7.7	1738	1	US-08-468-939-3
16	88.8	7.7	1738	2	US-08-406-855A-3
17	88.8	7.7	1738	2	US-08-722-190-3
18	88.8	7.7	1738	3	US-09-206-899-3
19	88.8	7.7	1738	3	US-08-244-354-3
20	88.8	7.7	1738	4	US-09-444-783-3
21	88.8	7.7	1738	4	US-09-688-415-3
22	88.8	7.7	1738	4	US-09-016-434-1402
23	88.8	7.7	1738	5	PCT-US95-04203-3
24	88.8	7.7	1738	5	PCT-US95-04203-3
25	88.8	7.7	2108	3	US-09-032-742-6
26	88.2	7.7	1134	1	US-08-087-772A-14
27	88.2	7.7	1227	1	US-07-916-901-1

28 88.2 7.7 1270 4 US-09-016-434-1184 Sequence 1184, Ap
29 88.2 7.7 3683 3 US-08-450-962-1 Sequence 1, Appli
30 88.2 7.7 3683 4 US-08-848-631-1 Sequence 1, Appli
31 85.6 7.4 1579 4 US-09-016-434-1243 Sequence 1243, Ap
32 84.6 7.3 1987 1 US-08-722-001-26 Sequence 26, Appl
33 84.6 7.3 1997 1 US-08-722-001-27 Sequence 27, Appl
34 84.6 7.3 2004 1 US-08-722-001-11 Sequence 11, Appl
35 83 7.2 1639 1 US-08-334-698-5 Sequence 5, Appli
36 83 7.2 1639 1 US-08-228-932-5 Sequence 5, Appli
37 83 7.2 1639 1 US-08-468-939-5 Sequence 5, Appli
38 83 7.2 1639 2 US-08-406-855A-5 Sequence 5, Appli
39 83 7.2 1639 2 US-08-722-190-5 Sequence 5, Appli
40 83 7.2 1639 3 US-08-244-354-5 Sequence 5, Appli
41 83 7.2 1639 3 US-09-206-899-5 Sequence 5, Appli
42 83 7.2 1639 4 US-09-444-783-5 Sequence 5, Appli
43 83 7.2 1639 4 US-09-688-415-5 Sequence 5, Appli
44 83 7.2 1639 4 US-09-444-783-5 Sequence 5, Appli
45 83 7.2 1639 5 PCT-US95-04203-5 Sequence 5, Appli

ALIGNMENTS

RESULT 1
US-08-031-538-1
; Sequence 1, Application US/08031538
; Patent No. 5968817
; GENERAL INFORMATION:
; APPLICANT: Sutcliffe, J Gregor
; APPLICANT: Erlander, Mark G
; APPLICANT: Lovenberg, Timothy W
; TITLE OF INVENTION: HUMAN SEROTONIN RECEPTORS, DNA ENCODING
; TITLE OF INVENTION: THE RECEPTORS, AND USES THEREOF
; NUMBER OF SEQUENCES: 73
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: The Scripps Research Institute, Office of
; ADDRESSEE: Patent Counsel
; STREET: 10666 No. 5968817th Torrey Pines Road, TPC 8
; CITY: La Jolla
; STATE: CA
; COUNTRY: USA
; ZIP: 92037
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/031,538
; FILING DATE: 19930315
; CLASSIFICATION: 800
; ATTORNEY/AGENT INFORMATION:
; NAME: Fitting, Thomas
; REGISTRATION NUMBER: 34,163
; REFERENCE/DOCKET NUMBER: TSR5099P
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 619-554-2937
; TELEFAX: 619-554-6312
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2226 base pairs
; TYPE: NUCLEIC ACID
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; HYPOTHEICAL: NO
; ANTI-SENSE: NO
US-08-031-538-1

Query Match 63.4%; Score 729.8; DB 2; Length 2226;
Best Local Similarity 79.6%; Pred. No. 1.3e-145;
Matches 911; Conservative 0; Mismatches 207; Indels 27; Gaps 3;

KW Pick's disease; Huntington's disease; Parkinson's disease; hypertension;
 KW atherosclerosis; myocardial infarction; gastritis; cirrhosis; cytostatic;
 KW osteoporosis; Crohn's disease; acquired immunodeficiency syndrome; AIDS;
 KW anaemia; asthma; rheumatoid arthritis; diabetes; obesity; drug screening;
 KW transgenic animal; allergy; gene therapy; hepatotropic; anticonvulsant;
 KW neurogenic; neuroprotective; cardiant; immunosuppressive; anorectic;
 KW virucide; receptor.

XX Homo sapiens.

XX Key Location/Qualifiers
 FT Peptide 1..65
 FT Domain 48..68 /label= Signal_peptide
 FT Protein 66..370 /note= "Transmembrane domain"
 FT Domain 298..316 /label= Human_mature_GCREC-15
 FT /note= "Transmembrane domain"

XX WO200210387-A2.

XX 07-FEB-2002.

XX 25-JUL-2001; 2001WO-US023433.

XX 27-JUL-2000; 2000US-0221478P.

XX 03-AUG-2000; 2000US-0223268P.

XX 21-AUG-2000; 2000US-0227054P.

XX 08-SEP-2000; 2000US-0231121P.

XX 13-SEP-2000; 2000US-0232243P.

XX 15-SEP-2000; 2000US-0232691P.

XX 22-SEP-2000; 2000US-0235146P.

XX (INCY-) INCYTE GENOMICS INC.

XX Thornton M, Patterson C, Lal P, Burford N, Yue H, Gandhi AR;

XX Elliot VS, Ramkumar J, Baughn MR, Kallick DA, Walla NK, Hafalia AJA;

XX Yao MG, Lu Y, Tribouley CM, Policky JL, Kearney L, Graul RC;

XX Warren BA, Lee EA, Ding L;

XX WPI; 2002-188744/24.

XX N-PSDB; AAD29681.

XX New human G-protein coupled receptor polypeptide for diagnosis,
 PT prevention and treatment of cell proliferative, neurological,
 PT cardiovascular, gastrointestinal, autoimmune/inflammatory, and metabolic
 disorders.

XX Claim 1; Page 133-134; 150pp; English.

XX The invention relates to novel human G-protein coupled receptors (GCREC)
 CC and their encoding polynucleotides. GCREC is useful as an immunogen for
 CC preparing monoclonal and polyclonal antibodies. GCREC is useful for
 CC diagnosing, treating and preventing a cell proliferative disorder (e.g.,
 CC hepatitis, psoriasis, cancer), a neurological disorder (e.g., epilepsy,
 CC Alzheimer's disease, Pick's disease, Huntington's disease, Parkinson's
 CC disease), a cardiovascular disorder (e.g., atherosclerosis, hypertension,
 CC myocardial infarction), gastrointestinal disorder (e.g., gastritis,
 CC cirrhosis, Crohn's disease), an autoimmune/ inflammatory disorder (e.g.,
 CC acquired immunodeficiency syndrome (AIDS), allergy, anaemia, asthma,
 CC rheumatoid arthritis), a metabolic disorder (e.g., diabetes, obesity,
 CC osteoporosis), and viral infections. GCREC is useful in a number of drug
 CC screening techniques, and to analyse the proteome of a tissue or cell
 CC type. GCREC is useful for creating knockin humanised animals or
 CC transgenic animals to model human diseases, in somatic or germline gene
 CC therapy, to generate a transcript image of a tissue or cell type, for
 CC detecting differences in the chromosomal location due to translocation,
 CC inversion, etc., among normal, carrier or affected individuals, and as
 CC hybridization probes for mapping naturally occurring genomic sequences.
 CC GCREC is useful in Southern or northern analysis, dot blot or other
 CC membrane-based technologies, in PCR technologies, in dipstick, pin,
 CC multi-format enzyme linked immunosorbant (ELISA)-like assays, and in

CC microarrays utilising fluids or tissues from patients to detect altered
 CC GCREC expression. The present sequence is human GCREC-15
 XX Sequence 370 AA;

Alignment Scores:
 Pred. No.: 1.86e-136 Length: 370
 Score: 1861.50 Matches: 370
 Percent Similarity: 97.63% Conservative: 0
 Best Local Similarity: 97.63% Mismatches: 0
 Query Match: 86.42% Indels: 9
 DB: 5 Gaps: 3

US-09-976-782-15 (1-1152) x AAEL8654 (1-370)

QY 5 ATGAGAGCGCTAGCCTTTCAGTGGCCACCGCGCGGTGGCTTGCCTGGAGCCGAG 64
 Db 1 MetGluAlaAlaSerLeuSerValAlaThrAlaGlyValAlaLeuAlaLeuGlyProGlu 20
 QY 65 ACCAGCAGCGGACCCCGGACCCCAAGCCCGAGAGGATACCTCGGTCGACCCCGAGCGGC 124
 Db 21 ThrSerSer-----GlyThrProSerProArgGlyLeuGlySerThrProSerGly 38
 QY 125 GCCGTCTCTGCGGCGCGAGGCGCCCTTCTCTCTTCAAGTCTCTGGTGGTGCAGCGTG 184
 Db 39 AlaValLeuProGlyArgGlyProProPheSerValPheThrValLeuValValThrLeu 58
 QY 185 CTAGTGTCTGTATCGCTGCGACCTTCTGTGTGGAACTCTGTGGTTCGGTCCACATCCG 244
 Db 59 LeuValLeuLeuIleAlaAlaThrPheLeuTrpAsnLeuLeuValProValThrIlePro 78
 QY 245 CGGTCTCGTGTCTTCCACCGCGTCCGCAATACCTTGTGTGGCTCGACGGCGCTCGGAC 304
 Db 79 ArgValArgAlaPheHisArgValProHisAsnLeuValAlaSerThrAlaValSerAsp 98
 QY 305 GAATAGTGGCAGCGCTGGCGATGCCACCGAGCCTGGCGAGTGTGTGACCGGGCGA 364
 Db 99 GluLeuValAlaAlaLeuAlaMetProProSerLeuAlaSerGluLeuSerThrGlyArg 118
 QY 365 CGTGGTGTGGCGCGGAGCGCTGTGCGACGTGTGGATCTCTTCCGACGCGCGAGCGCTGT 424
 Db 119 ArgArgLeuLeuGlyArgSerLeuCysHisValTrpIleSerPheAspAla----- 135
 QY 425 CTGTGTCTGCCCGCGCGCTCGGAAAGTGGCGGCATCGCCCTGGCGCGAGCGGGCC 484
 Db 136 LeuCysProAlaGlyLeuGlyAsnValAlaAlaIleAlaLeuGlyArgAspGlyAla 155
 QY 485 ATCACAGCGCACCTTGACGACACGCTGCGCACCCCGAGCGCGCTCGTTCCTCATGATC 544
 Db 156 IleThrArgHisLeuGlnHisThrLeuArgThrArgSerArgAlaSerLeuLeuMetIle 175
 QY 545 GCGCTCCCGCGGTGCGTGGCGCTCATGCGCTCCGCGCGGTGTCTTTGGCGGGCG 604
 Db 176 AlaLeuAlaArgValProSerAlaLeuIleAlaLeuAlaProLeuLeuPheGlyArgGly 195
 QY 605 GAGGTGCGCAGCGCTCGGCTCCGCGCTGCGAGTGGAGCGGGAACCTCTATGCGCGC 664
 Db 196 GluValCysAspAlaArgLeuGlnArgCysGlnValSerArgGluProSerTyrAlaAla 215
 QY 665 TTCTCCACCGCGCGCGCTTCCACCTGCGCGTGGCGGTGGCGGTGTCTTACCGGAAG 724
 Db 216 PheSerThrArgGlyAlaPheHisLeuProLeuGlyValValProPheValTyrArgLys 235
 QY 725 ATCTACAGCGCGGCAAGTTGCTTTCGCGCGCGCGAGAGTGTGTGCGCGTCCGCG 784
 Db 236 IleTyrGluAlaAlaLysPheArgPheGlyArgArgArgAlaValLeuProLeuPro 255
 QY 785 GCCACCATGAGGTGAGGTCCAGGTAAAGAGACACCTCATGAGTGGTGAAGTGTTC 844
 Db 256 AlaThrMetGln-----ValLysGluAlaProAspGluAlaGluValValPhe 271
 QY 845 ACGGCACATTGCAAGCAACGGTGTCTTCCAGGTGAGCGGGGACTCTCTGGCGGGAGCAG 904

127 ATCTACCTTGTGGCTTCTGGTGGCGGACGTTGGCTGGAACCTGCTGGTGTG 186
Db
233 GTACACATCCGCGGTCGGTCCACCGCTGCGCATACTTGGTGGCTTCAGC 292
Qy
187 GCGACATCTCCGTGACGACCTTCCACCGCTGCCACACCTGGTGGCATCATG 246
Db
293 GCGCTCTCGGACGAACCTAGTGGGAGCGCTGGCGATGACACGAGCTGGCGAGTGTG 352
Qy
247 GCGCTCTCGATGTCCTGGTGGCGCGCTGGTATGCGCTGAGCTTGGTGCAGAGTGT 306
Db
353 TCGACCGGCGGAGTCCGCTGCTGGGCGGAGCCTGTGCCACGCTGTGGATCTTCTTCGAC 412
Qy
307 T--CGGGCGCGCTCGGACGCTAGTGGGAGGCTGTGCCAGCTTGGATCGGCTGCGAC 363
Db
413 GCGGAGCTGTGTGTGCTGCGCGCGCTCGGAGCTGGGCGCATCGCCCTGGC 472
Qy
364 G-----TGCTTGTGTCAGCGGCGAGCATCTGGAACGTGAGCGGCATAGCAGTGCAC 414
Db
473 GCGGAGCGGCGCATCACACGCGCCTCGACACACGCTGGCGACCGCGCGCGCTCG 532
Qy
415 CGCTACTGTGTCATCAGCGGCCATGGAATACACGCTCCGACCGCGAGTGGTCTCC 474
Db
533 TTGCTATGATCGGCTTCCGCGCGGTGCGTGGCGCTCATGCGCTTCGCGCGCTGCTC 592
Qy
475 AAGCTCATGATCGGCTCACCTGGGCACTCNCACCTGTCTCTCTGGCGCGCTGCT 534
Db
593 TTTGGCGGCGGAGGTGTCGAGCTCGGCTCCAGCGCTGCCAGTGGAGCGGGAACCC 652
Qy
535 TTTGGCTGGGAGAGAGTACTCTGAGGCGAGCGAGGAGTGCACAGTAAAGCGGAGCCT 594
Db
653 TCCTATGCGCGCTTCTCACCGCGCGCTTCCACCTGCGCTGGCGTGGTGGCTGCTG 712
Qy
595 TCCTAGCGCGTGTCTCACCGTAGGCGCTTCTACCTGCGCTCTGTGTGGTCTTC 654
Db
713 GTCTACCGGAAGTCTACGAGGCGGCAAGTTTGGTTTCGCGCGCGCGCGAGAGCTGTG 772
Qy
655 GTGTACTGGAAGATCTACAAGGCTACCAAGTTCGCGTGGCTCCAGGAAGACCAATAGC 714
Db
773 CTGCGGTTCGCGGCGCACCATGAGTGGTCCAGTCCAGTAAAGGAGCACCTGATGAGCT 832
Qy
715 GT-----CTCACCCTATCCGAAGCTGTGGAGGTGAAGGACTGTGCCAACAGCCC 765
Db
833 GAAGTGTGTTCACGCGCACATTCGAAGCAAGCTGTCTTCCAGGTGAGCGGAGCTCC 892
Qy
766 CAGATGTTTACGGTCC---GCCAGCCACCGTACCTTCCAGCCAGAGGAGGACAG 822
Db
893 TGGCGGAGCAGAGGAGGAGCGAGCCATGATGTTGGGAAATCTGATTGGCGTGT 952
Qy
823 TGTGGGAGCAGAGGAGCAGCGCGCGCTCATGTTGGGCTCTCTCATTTGGCGTGTTC 882
Db
953 GTGCTGTGCTGGATCCCTTCTTCTGACCGAAGTATGACCGCACTCTGTGCTGAGC 1012
Qy
883 GTGCTGTGCTGGATCCCTTCTTCTCAGCGAGCTCATGATCCCTCTCTCTCTGTGAC 942
Db
1013 CTGCGCGGCTATCGGAAAGCATATTTCTGTGGCTTGGCTACTCCAAATCTTCTTCAAC 1072
Qy
943 ATCCCGGCTATCGGAAAGCATCTTCTGTGGCTTGGCTACTCCAACTCTTCTTTAC 1002
Db
1073 CCCTGATTTACAGCTTTTAAAGAACTTAAACAATGCTTCAAGAGCCTCTTTTACT 1132
Qy
1003 CCGCTGATCTATACGGCTTTCAAGAACTTAAACAAGAGCTTCAAGAGCCTCTTTTCT 1062
Db
1133 AAGCA 1137
Qy
1063 AGGCA 1067
Db

RESULT 3

US-08-356-405-1
; Sequence 1, Application US/08356405
; Patent No. 5807691
; GENERAL INFORMATION:

APPLICANT: Amlaiky, No. 5807691rdine
APPLICANT: Boschert, Ursula
APPLICANT: Hen, Rene
APPLICANT: Plasseat, Jean-Luc
TITLE OF INVENTION: Polypeptides Having Serotonin Receptor
TITLE OF INVENTION: Activity (SHTSA), Nucleic Acids Coding for These
TITLE OF INVENTION: Polypeptides and Uses Thereof
NUMBER OF SEQUENCES: 13
CORRESPONDENCE ADDRESS:
ADDRESSEE: Rhone-Poulenc Rorer Inc.
STREET: 500 Arcola Road, 3C43
CITY: Collegeville
STATE: PA
COUNTRY: USA
ZIP: 19426-0107
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/356,405
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: WO PCT/FR93/00650
FILING DATE: 29-JUN-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: FR 92/08081
FILING DATE: 01-JUL-1992
ATTORNEY/AGENT INFORMATION:
NAME: Smith, Julie K
REGISTRATION NUMBER: 38,619
REFERENCE/DOCKET NUMBER: EX92004-US
TELECOMMUNICATION INFORMATION:
TELEPHONE: (610)454-3839
TELEFAX: (610)454-3808
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 1686 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
FEATURE:
NAME/KEY: CDS
LOCATION: 509..1582
US-08-356-405-1

Query Match 36.9%; Score 424.6; DB 1; Length 1686;
Best Local Similarity 67.3%; Pred. No. 4.1e-81;
Matches 684; Conservative 0; Mismatches 309; Indels 24; Gaps 5;
Qy 128 GTCTGCGGCGGAGGCGCGCTTCTGTCTTTCACGCTCTGTGGTGGAGCTGCTA 187
Db 590 GTCTGCGGCTAGTCGGCTTTTCTCAGCTTTCGAGTGTAGTCTTGTG 649
Qy 188 GTCTGCTGATCGCTGCGCACTTTCCTGTGGAACCTGTGTGGTTCGGTTCACATCCCGCG 247
Db 650 GGCTTTCTAGCTCGGCGCACATTCACCTTGGAACTGTGTGTGGTGTGGCTACCTCAAG 709
Qy 248 GTCCGCTGCTTCCACCGCTGCGCATTAAGTGTGGCTCGAGCGGCTCGGAGCA 307
Db 710 GTACGACCTTCCACCGAGTACCAACAACTGTAGCTTCCATGGCCATCTCGGATGT 769
Qy 308 CTAGTGGCAGCGCTGGCGATGCCACGAGCTCGCGAGTGTGAGCTGTGCGCGCGACGT 367
Db 770 CTAGTGGCTGTGCTGTTATGCGCACTGAGCTGTGATGAGCTGT---CTGGCGCGCG 826
Qy 368 CGGCTGCTGGCGGAGCCTGTGCCACAGTGTGATCTTCTTCCAGCGCGGAGCCTGTCTG 427
Db 827 TGGCAGCTGGCGCGGCTCTATGCCAGCTGTGATCGCATGTGACG-----TGCTC 877

Qy	428	TGCTGCCCCCGGCTCTGGGAACGTGGCGGCCAATCGCCCTGGGCGCGGACGGGGCAATC	487
Db	878	TGCTGTACTGCCAGCATCTGGAATGTACACAGCAATAGCACTGGACCGCTACTGTCAAATC	937
Qy	488	ACACGGCACCTGGAGCACAGCTCGCACACCGGAGCGCGCTGCTGTGCTCATGATCGCG	547
Db	938	ACGGCCACCTGGAGTACACACTCGGTACCGCAAGCGTGTCTCCAAATGTGATGATCTCTG	997
Qy	548	CTCGCCCGGGTTCGCGTCTGGCGCTCATCGCCCTCGCGCGCTGCTCTTTGGCCGGGCGAG	607
Db	998	CTCACCTGGGCACCTCTCCACTGTTCATCTCTCTGCTCCACTGTCTATTTTGCTGGGAGAG	1057
Qy	608	GTGTGCGACGCTCGGCTTCGAGCGTGCACAGTGAGCGGAAACCTCTCATGTCGCGCTTC	667
Db	1058	ACTTATTCTGAGCCCGAGTGAGGAATGCCAGTCAGTCGCGAGCCTTCTCTACACCGTGTTC	1117
Qy	668	TCCACCCGCGCGCTTCCACCTCCCGCTTGGCGTGTGCGCTGTGCTTGTCTACCGGAAGATC	727
Db	1118	TCCACCGTGGTGCTTCTACCTCCCGCTGTGGCTGTGTCTTTGTGTACTTGGAAAAAT	1177
Qy	728	TACAGAGCGGCGAATTTGCTTTTGGCGCGCGCGGAGAGCTGTGTCCGTTGCCGGCC	787
Db	1178	TACAGGCGGCGAAATTCGCGATGGGCTCCAG-----GAAGACTAACAGCGTCTCCCCCG	1232
Qy	788	ACCATGCAAGTGAGCTCCAAAGTAAAGGAAGCACCTGATGAGGCTGAAGTGGTGTTCACG	847
Db	1233	TACCCGAAGCTGTGG-----AGGTGAAGAATGCTACACAAATCCCCAGATGGTGTTCACG	1288
Qy	848	GCAATTTGCAAGCAACGGTGTCTCTCAGGTGAGCGGGACCTCTGGCGGAGACAGAAG	907
Db	1289	GCCC---GCCATGCCACCGTCACCTTCCAGACAGAAGGGGATACGTGGAGGGAGCAGAAG	1345
Qy	908	GAGAGCGAGCAGCATCATCGTGGGAATTCGTATTGGCGTGTGTTGTGCTGTGTGGATC	967
Db	1346	GAGCAAGGCGAGCCCTCATGTGGGCATCTCTATCGAGGTGTTGTGCTCTGCTGTTCT	1405
Qy	968	CCCTTCTTCTGACGGAACATCATCAGCCCCACTGTGTGCCTGCAGCCTGCCCCCCCAATCGG	1027
Db	1406	CTTTTCTTCGTACAGAGCTCATCAGTCCCTGTGTTTCTCTGGGACGTCCCTGTCCATCTGG	1465
Qy	1028	AAAAGCATATTTCTGTGGCTTGGCTACTCCAAATTTCTTCTTCAACCCCTTGATTACACA	1087
Db	1466	AAGAGCATTTCTCTGTGGTATTTCTAAATTTCTTCTTCAACCCCATCATCTACACA	1525
Qy	1088	GCTTTTAAAGAACTACAACAATGCCTTCAAGAGCCTCTTTTACTAAGACAGATGA	1144
Db	1526	GCATTTCAACAGAGCTACAGCAGTGTCTTCAAGTCTTCTTCTCTCAAGCAACAATGA	1582

RESULT 4

US-08-031-538-3
Sequence 3, Application US/08031538
Patent No. 5968817
GENERAL INFORMATION:
APPLICANT: Sutcliffe, J Gregor
APPLICANT: Brlander, Mark G
APPLICANT: Lovenberg, Timothy W
TITLE OF INVENTION: HUMAN SEROTONIN RECEPTORS, DNA ENCODING
TITLE OF INVENTION: THE RECEPTORS, AND USES THEREOF
NUMBER OF SEQUENCES: 73
CORRESPONDENCE ADDRESS:
ADDRESS: The Scripps Research Institute, Office of
ADDRESS: Patent Counsel
STREET: 10666 No. 5968817th Torrey Pines Road, TPC 8
CITY: La Jolla
STATE: CA
COUNTRY: USA
ZIP: 92037
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25


```

; APPLICATION NUMBER:
; FILING DATE:
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Zeller, Karen J.
; REGISTRATION NUMBER: 37,071
; REFERENCE/DOCKET NUMBER: PA-0002 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (650) 855-0555
; TELEFAX: (650) 845-4166
; INFORMATION FOR SEQ ID NO: 1077:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1984 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; LIBRARY: GENBANK
; CLONE: g1162923
;
US-09-016-434-1077

Query Match      8.3%; Score 95.2; DB 4; Length 1984;
Best Local Similarity 47.1%; Pred. No. 1.9e-11;
Matches 489; Conservative 0; Mismatches 513; Indels 36; Gaps 5;

QY 114 CCCGAGCGGCGCGTCTCGCGGGCCGAGGGCGCCCTTCTCTGCTTTCACGGTCCCTGG 173
DB 493 CCAATAGCACCCCGGCTTGGGGGGCAGGGCGCGCTCGGCCCGGGGCGAGCGGCTGGG 552

QY 174 TGGTGACGCTGTAGTGTGCTGATCGCTGCACCTTTCTGTGGAACCTGTGGTTCCGG 233
DB 553 TGGCGGCGCGCTGTGCGTGTGTCATCGCGCTGACGGGGCGGCCAACTCGCTGATCG 612

QY 234 TCACATACCCGGGGTCCGCTTCCACCGGTGCCGCAAACTTGTGTGGCTCGACGG 293
DB 613 CGCTCATCTGCACTCAGCGCGGCTGCGCAACACGCTCCAACTTCTCTGCTGTGCTCT 672

QY 294 CCGTCTCGAGCAACTAGTGTGCGAGCGCTGGGGATGCCACCGAGCTGCGAGTGAAGTGT 353
DB 673 TCAGCTGCACTGATGTGGGGCTGGTGTGATGCCCGCGCCNAGCTGAACGCGCTGT 732

QY 354 CGACGGGCGAGCTCGGCTGTGGCGCGGAGCTGTGCAAGTGTGATCTCTTCGACG 413
DB 733 ACG-----GGCGTGGTGTGGCGCGGCTCTGCTGTCTGGACCGCTTCGACG 786

QY 414 CGGAGCCCTGTGTGCTGCCCGCGCGCTCGGGAAAGTGGCGCCATGCCCTGGGCC 473
DB 787 -----TGATGTGTGCGAGCGGCTCCATCTCAAGCTTGTGCTCATCAGCCTGGACC 837

QY 474 GCGAGGGGCGCATCACACGGCACCTGCGAGCACAGCTGCGCACCGCGCGCCTCGT 533
DB 838 GCTACCTGCTATCTCTGCGCGTGGCTTACAGCTGGGATGAGCCCTCGGTGGCC 897

QY 534 TGCTCATGATCGGCTCGCCCGGGTGGCTGCGGCTCATCGCCCTCGCGCGGCTGCTCT 593
DB 898 TGGCCCTAGTCTTGGGCGCTTGGAGCTTGGCGCTCTCGCTCTCTTCTGCGCCCTGTGC 957

QY 594 TTGGCGGGGCGAGTGTGCGACGCTCGGCTCCAGCGCTGCCAGTGTGAGCGGGAACTT 653
DB 958 TGGGCTGCGAGAGTGGGGCCACGACGCGC-----CACCGTCCCTCG 1000

QY 654 CCTATGCGGCTTCTCCACCGCGCGCTTCCACTTGGCGTGGCGTGGTGGCGTTTG 713
DB 1001 CAGTGGCGGCTGTGGCGAGCTGCTTGTGCTTGTGGGTGGGCGCTCACCTTCTT 1060

QY 714 TCTACCGAAGATCTACAGAGCGGCAAGTTTCGTCGCGCGCGCGGAGAGCTGTGC 773
DB 1061 CTTGCCCTCGGGTGCATATGCTTCACTTCTGAGGATCTCTGTAGTGTGCCCGC-AAAG 1119

QY 774 TGCCGTTGCCGCGACCATGAGGTGAGTCCAAAGTAAGAAAGCACTGATGAGGCTG 833
DB 1120 AGGCGGTGAGGTGGCTTCCCTCAACACCGGATGGCCAGTCCAGGCTCGGAGAGCTGCTG 1179
```

```

QY 834 AAGTGTGTTCACGGCACATTTGCAAGCAACGGTGTCTTCCAGGTGAGCGGGACTCTCT 893
DB 1180 AGTGTCCAGGACCCCAAGCCAGGGGTGGAGTCTGCTGACAGAGGCGTCTAGCCAGA 1239

QY 894 GGGGGAGCAGAGGAGAGGCGAGCCATGATGGTGGGAATCTGATGGCGTGTGTTG 953
DB 1240 AGCAGCAGGAGAGGCGCTGAAGCCAGCCCTGACGCTGGGCATCTGCTGGGCAATGTTCT 1299

QY 954 TGCTGTGTGATCCCTCTTCTTCTGACGGAACCTATCAGCCCACTCTGTGCTGCAGCC 1013
DB 1300 TTGTGACCTGGTGGCCCTTCTTGTGGCCAACTAGTCCAGGCGTGTGCGACTGATCT 1359

QY 1014 TGCCCCCATCTGAAAAGCATATTTCTGTGGCTTGGCTACTCCAAATCTTTTTCACCC 1073
DB 1360 CCCAGGCGCTTCTGA---TGTCTCATATGGCTGGGTACTGTAAACAGCACCATGAACC 1416

QY 1074 CCCTGATTACAGCTTTTAAACAAGAACTAACAATGCTTCAAGAGCCTCTTTACTA 1133
DB 1417 CCATCATCTACCCACTCTTTCATGCGGGACTTCAAGCGGGCGCTGGGCAAGTTCCTGCCAT 1476

QY 1134 AGCAGATGAACACAGG 1151
DB 1477 GTCCACGCTGTCCCGGG 1494

RESULT 7
US-09-016-434-1197
; Sequence 1197, Application US/09016434
; Patent No. 6500938
; GENERAL INFORMATION:
; APPLICANT: Janice Au-Young
; APPLICANT: Jeffrey J. Seilhamer
; TITLE OF INVENTION: COMPOSITION FOR THE DETECTION OF SIGNALING
; TITLE OF INVENTION: PATHWAY GENE EXPRESSION
; NUMBER OF SEQUENCES: 1490
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: INCYTE PHARMACEUTICALS, INC.
; STREET: 3174 PORTER DRIVE
; CITY: PALO ALTO
; STATE: CALIFORNIA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Word Perfect 6.1 for Windows/MS-DOS 6.2
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/016,434
; FILING DATE: HEREWITH
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Zeller, Karen J.
; REGISTRATION NUMBER: 37,071
; REFERENCE/DOCKET NUMBER: PA-0002 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (650) 855-0555
; TELEFAX: (650) 845-4166
; INFORMATION FOR SEQ ID NO: 1197:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1756 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; LIBRARY: GENBANK
; CLONE: g181431
;
US-09-016-434-1197
```


Db 710 CAAGAACCTAGAGCGAGGAGTCAAGAGAGATGTCCAACTCCAGAGAGCTGACCCTG-- 767
QY 811 AAAGGAAGCACTGATGAGGCTGAAGTGGTGTTCACGGCAGCATTCGAAAGCAAGGTTGTC 870
Db 768 --AGGATCCATTCCAAAGAACTTTACAGAGACACCTTTAGCAGTACCAAGGCCAAGGGCC 825
QY 871 CTTCCAGGTAGCGGGGACTCCTGGCG-----GGAGCAGAGGAGGCGAGCAG 920
Db 826 ACAACCCAGGAGTTCATAGCTGTCAAACTTTTAAAGTCTCCAGGGAAGAAAGACAG 885
QY 921 CCATGATGGTGGGAATTCATGATGGCGCTGTTTGGCTGTGTGATCCCTCTTCTCTGTA 980
Db 886 CTAAGAGCTTGGCAATGTGGTGGTATGTTTCATCTTGTGCTGGTACCTCTTCTATCG 945
QY 981 CGGAACATCATAGCCCACTCTGTGCCCTGTCAGCCTGCCCC---CATCTGGAAGAGCATAT 1037
Db 946 CTCCTACCGCTTGGCTCTCTGTTCTCCACCTGAAGCCCGCCGAGCGCGTGTTCAGGTGG 1005
QY 1038 TTCTGTGGCTTGGCTACTCCAAATCTTTCTTCAACCCCTGATTTACAGCTTTTAACA 1097
Db 1006 TGTCTGTGGCTGGGCTACTTCAACAGCTGCGCTCAACCCCATCATCTACCCATGCTCCAGCA 1065
QY 1098 AGAACTACAAATGCTTCAAGAGCCTCTTT 1129
Db 1066 AGGAGTTCAAGCGGCTTTCATGGTATCCTT 1097

RESULT 12

US-08-351-473B-7
; Sequence 7, Application US/08351473B
; Patent No. 5656440
; GENERAL INFORMATION:
; APPLICANT: LENZEN, GERLINDA
; APPLICANT: KAPOOR, ARCHANA
; TITLE OF INVENTION: NUCLEOTIDE SEQUENCES CODING FOR THE
; TITLE OF INVENTION: BOVINE BETA3-ADRENERGIC RECEPTOR AND THEIR APPLICATIONS
; NUMBER OF SEQUENCES: 9
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: OBLON, SPIVAK, MCLELLAND, MAIER & NEUSTADT
; STREET: 1755 S. JEFFERSON DAVIS HIGHWAY, SUITE 400
; CITY: ARLINGTON
; STATE: VIRGINIA
; COUNTRY: USA
; ZIP: 22202
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/351.473B
; FILING DATE: 21-FEB-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 93 04670
; FILING DATE: 21-APR-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/FR94/00447
; FILING DATE: 21-APR-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: OBLON, NORMAN F.
; REGISTRATION NUMBER: 24,618
; REFERENCE/DOCKET NUMBER: 6639-001-0X PCT
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703) 413-3000
; TELEFAX: (703) 413-2220
; TELEX: 248855 OPAT UR
; INFORMATION FOR SEQ ID NO: 7:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1227 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single

; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
US-08-351-473B-7
Query Match 7.8%; Score 89.8; DB 1; Length 1227;
Best Local Similarity 48.6%; Pred. No. 2.4e-10;
Matches 460; Conservative 0; Mismatches 442; Indels 45; Gaps 6;
QY 165 CGGTCTCTGGTGGTGAAGCTGCTAGTGTGCTGATCGTGCACCTTCTCTGTGGAACCTGC 224
Db 110 CGGCCCTTAGCCGGGGCCCTGCTGGCGCTGGCGGTGTGGCCACCGTGGGAGGCAACTGC 169
QY 225 TGGTTCGGGTCAACATCCCGCGGCTCCGTCCTCCACCGCGTGCACATAAATCTTGGTGG 284
Db 170 TGGTCATCTGGCCATCGCTCGACTCCGAGACTCCAGACCATGACCAACGTGTTCGTGA 229
QY 285 CTTTCGAGCGCGCTCTCGAGCAACTAGTGGCAGGCTGGCGATGTCACACGAGCCTGCGGA 344
Db 230 CTTTCGCTGGCGGAGCGACCTGGTGTATGGACCTCTCTGGTGGTGGCGCC-----GGCGG 283
QY 345 GTGAGCTGTGACCGGGCGAGCTGGCTGCTGGCGCGAGCCTGTGCGACAGTGTGATCT 404
Db 284 CCACCTTGGCGCTGACATGGCCACTTGGCCGTTGGCGCCACTGGCTGCGAGCTGTGACCT 343
QY 405 CTTTCGAGCGCGGAGCCTGTGTGTGTCGCCCGCGGCTCGGGAAGCTGGCGGCCATCG 464
Db 344 CGGTGGAGTG-----CTGTGTGACCGCGCAGCATCGAAACCTGTGGCCCTCGG 394
QY 465 CCTTGGCGCGGAGCGGCCATCACCGGCACTGACAGCACCTGGGCAACCGCGAGCC 524
Db 395 CCGTGGACCGCTACTGGCTGTGACCAACCGCTGCGTTAGCGGGCAGCTGGTCAACCAAGC 454
QY 525 GCGCTCGTGTCTCATGATCGCGCTCGCGCGGCTGGCGCTCATCGCCCTCGCGC 584
Db 455 GCTGCGCCCGGACAGCTGTGTCTGTGTGGGTGCTGTGCGCGCGGTGTGTTTGGC 514
QY 585 C--GCTGCTCTTTGGCGGGCGAGGTGTGCGACGCTCGGCTCCAGCGCTGCCAGTGA 641
Db 515 CCATCATGAGCAGTGTGTGGCGGTAGGGGCCGACGCGGAGCGCAGCGCTGCCACTCA 574
QY 642 GCGGGGAACCTCTATGCGCCCT-----TCTCCACCGCGCGG 680
Db 575 ACCCGCTGCTGTGCGCTTGGCTTCCAAACATGCGCCTACGTGCTGTCTCTCTCCCTCT 634
QY 681 CTTTCACCTGCGCTTGGGTGTGGTGTGCTTGTCTACCGGAAGATCTACGAGCGGCA 740
Db 635 CTTTCTACCTTCTCTCTGTAIGCTCTTCGTCTACGCGCGGTTTCTGTTGGCTA 694
QY 741 AGTTTCGTTTCGCGCGCGCGAGAGCTGTGCTGCGGTTGCGCGGCCACCATGCAAGTGA 800
Db 695 CGGCCAGCTGCGTTTGTGCGGGGAGCTGGGCGCTTTCG--CCGAGGAGTCTC 751
QY 801 GGTCCAAAGGTAAAGGAAGCCTGATGAGGTGAAGTGGTGTTCACGGCACAATTGCAAG 860
Db 752 GCGCGGCGCGCTGCGCTCTCTGCGCCCGCGCGGTGGGACGCTGCGCTCGCGCCGAAG 811
QY 861 CAACGCTGTCTTCCAGGTGAGCGGGAGCTCTGCGCGGAGCAGAGAGAGGAGGAGCAG 920
Db 812 GGGTGGCGCGCTGCGCGCGCGCGCGCGCTCTCTGCGCTCTCGGGAACACCGGGGCC 871
QY 921 CCATGATGGTGGGAATTCATGATGGCGTGTGTGTGCTGTGATCCCTTCTTCTCTGA 980
Db 872 TGTGCACTTGGGTCTCATCATGGGCACTTCACTCTCTGCTGTGTTGCTTCTTCTGG 931
QY 981 CGGAACATCATAGCCCACTCTGTGCTGCGAGCCTGCGCCCGCATCTGGAAGAGCATTTTC 1040
Db 932 CCAACGCTGCTGCGCGCTTGGGGGGCGCTCTCTAGTCCCGGGCGCGCTTCTCTTCTGCC 991
QY 1041 TG---TGGCTTGGCTACTCCAAATCTTCTTCAACCCCTGATTTAC 1084
Db 992 TGAACCTGGCTAGGTTATGCCAAATCTGCTTCAACCCCGCTCATCTAC 1038


```

; STRANDEDNESS: single
; TOPOLOGY: unknown
; MOLECULE TYPE: DNA (genomic)
; HYPOTHETICAL: N
; ANTI-SENSE: N
; FEATURE:
;   NAME/KEY: CDS
;   LOCATION: 124..1683
;   OTHER INFORMATION:
US-08-228-932-3

Query Match
Best Local Similarity 7.7%; Score 88.8; DB 1; Length 1738;
Matches 366; Conservative 0; Mismatches 362; Indels 20; Gaps 4;

QY 394 CGTGTGGATCTCTTCGAGCGGAGCGCTGTCTGTGCTGCCCCCGGCGCTCGGGAACGT 453
Db 471 CTTCTGTGACATCTCGGCGAGCGTGGATGCTCTGTGCTGCACAGCGTCCATTCTCAGCCT 530

QY 454 GCGCGCCATCGCCCTGGGCGGAGCGGCCATCACCGGACCTGCAGCACACGCTGCG 513
Db 531 GTGCGCCATCTCCATCGATCGTACTCATCGGGTGGCTACTCTCTGAGTATCCACGCT 590

QY 514 CACCGGACCGCGCTCTGTCTCATGATCGCGTTCGCGGGTGCCTGCGGCGCTCAT 573
Db 591 GGTCAACCGGAGGAGGCGCATCTTGGCGCTGCTCAGTGTCTGGTCTTGTCCACGCTCAT 650

QY 574 CGCCCTCGCGCGCTCTTGGCGGCGGAGGTGTGCGAGCGTTCGCGTCCAGCGGTG 633
Db 651 CTCATCGGCGCTCTCTTGGTGAAGGAGCGGCGCACCCACGAT---GACAAGGAGTG 707

QY 634 CCAGGTGAGCGGGAACCTCTATGCGCGCTTCTCCACCGCGCGGCTTCCACTGCC 693
Db 708 CGGGTCAACGAGAACCTCTTATGCGCCTCTTCTCTCTCTGCGCTCTTCTACATCCC 767

QY 694 GCTTGGCGTGTGCGTGTGTACCGGAAGATCTACGAGGCGGCGCAAGTGTTCGG 753
Db 768 TCTGGCGTGTCTTAGTCTAGTCTAGTCTAGTCTAGTCTAGTCTAGTCTAGTCTAG 827

QY 754 CCGCGCGGAGAGCTGTGCTGCGCTTTCGCGGCGGAGGTGTGCGAGCGTTCAGGTCAA 813
Db 828 GAACCTAGAGGAGGATCATGAAGGAGATGTCCAACTCCAAAGGAGTGAACCTG---A 883

QY 814 GGAAGCACCTGATGAGGCTGAAGTGGTGTTCACGGCACATTCGCAAGCAACGGTGTCTT 873
Db 884 GGATCAATTCGAAGAACTTTCACGAGGACACCTTAGCAGTACCAAGGCGGAGGCGCACA 943

QY 874 CCAGGTGAGCGGGACTCTCTGGCG-----GGAGCAGAGGAGGAGGAGGAGGAGGAG 923
Db 944 ACCCGAGAGTTCATAGCTGTCAAACTTTTAAAGTTCCTCAGGGGAAAGAAAGCAGCTA 1003

QY 924 TGATGTGGGAATTCGATGCGGTGTTGTGCTGTGCTGATCCCTTCTTCTCTGAGCGG 983
Db 1004 AGAGTGGGCAATGTGTCGATGTTTCATCTTGTGCTGCTACCTTCTTCTTCTTCTGCTC 1063

QY 984 AACTCATCAGCCCACTCTGTGCTGCGCTGCGCTGCCCC---CATCTGGAAGAGCATATTC 1040
Db 1064 TACCGTTGGCTCTCTTGTCTCCACCTTGAAGGCGGCGGCGGCTGTTCAGGTGGTGT 1123

QY 1041 TGTGGCTGGCTACTCCAAATCTTCTTCAACCGGCTGATTATACAGCTTTTAAACAAGA 1100
Db 1124 TCTGGCTGGCTACTTCAACAGCTGCTCAACCGGCTGCTCAACCGGCTGCTCAACCGGCT 1183

QY 1101 ACTCAACAATGCTTCAAGAGCTCTT 1128
Db 1184 AGTTCAAGCGGCTTTCGTGGCATCCT 1211

```

RESULT 15

```

US-08-468-939-3
; Sequence 3, Application US/08468939
; Patent No. 5714381
; GENERAL INFORMATION:

```

```

; APPLICANT: Jonathan A. Bard et al.
; TITLE OF INVENTION: DNA Encoding Human Alpha 1 Adrenergic
; MOLECULE TYPE: DNA (genomic)
; HYPOTHETICAL: N
; ANTI-SENSE: N
; FEATURE:
;   NAME/KEY: CDS
;   LOCATION: 124..1683
;   OTHER INFORMATION:
US-08-228-932-3

Query Match
Best Local Similarity 7.7%; Score 88.8; DB 1; Length 1738;
Matches 366; Conservative 0; Mismatches 362; Indels 20; Gaps 4;

QY 394 CGTGTGGATCTCTTCGAGCGGAGCGCTGTCTGTGCTGCCCCCGGCGCTCGGGAACGT 453
Db 471 CTTCTGTGACATCTCGGCGAGCGTGGATGCTCTGTGCTGCACAGCGTCCATTCTCAGCCT 530

QY 454 GCGCGCCATCGCCCTGGGCGGAGCGGCCATCACCGGACCTGCAGCACACGCTGCG 513
Db 531 GTGCGCCATCTCCATCGATCGTACTCATCGGGTGGCTACTCTCTGAGTATCCACGCT 590

QY 514 CACCGGACCGCGCTCTGTCTCATGATCGCGTTCGCGGGTGCCTGCGGCGCTCAT 573
Db 591 GGTCAACCGGAGGAGGCGCATCTTGGCGCTGCTCAGTGTCTGGTCTTGTCCACGCTCAT 650

QY 574 CGCCCTCGCGCGCTCTTGGCGGCGGAGGTGTGCGAGCGTTCGCGTCCAGCGGTG 633
Db 651 CTCATCGGCGCTCTCTTGGTGAAGGAGCGGCGCACCCACGAT---GACAAGGAGTG 707

QY 634 CCAGGTGAGCGGGAACCTCTATGCGCGCTTCTCCACCGCGCGGCTTCCACTGCC 693
Db 708 CGGGTCAACGAGAACCTCTTATGCGCCTCTTCTCTCTCTGCGCTCTTCTACATCCC 767

QY 694 GCTTGGCGTGTGCGTGTGTACCGGAAGATCTACGAGGCGGCGCAAGTGTTCGG 753
Db 768 TCTGGCGTGTCTTAGTCTAGTCTAGTCTAGTCTAGTCTAGTCTAGTCTAGTCTAG 827

QY 754 CCGCGCGGAGAGCTGTGCTGCGCTTTCGCGGCGGAGGTGTGCGAGCGTTCAGGTCAA 813
Db 828 GAACCTAGAGGAGGATCATGAAGGAGATGTCCAACTCCAAAGGAGTGAACCTG---A 883

```

Query Match

```

Best Local Similarity 7.7%; Score 88.8; DB 1; Length 1738;
Matches 366; Conservative 0; Mismatches 362; Indels 20; Gaps 4;

```

```

QY 394 CGTGTGGATCTCTTCGAGCGGAGCGCTGTCTGTGCTGCCCCCGGCGCTCGGGAACGT 453
Db 471 CTTCTGTGACATCTCGGCGAGCGTGGATGCTCTGTGCTGCACAGCGTCCATTCTCAGCCT 530

QY 454 GCGCGCCATCGCCCTGGGCGGAGCGGCCATCACCGGACCTGCAGCACACGCTGCG 513
Db 531 GTGCGCCATCTCCATCGATCGTACTCATCGGGTGGCTACTCTCTGAGTATCCACGCT 590

QY 514 CACCGGACCGCGCTCTGTCTCATGATCGCGTTCGCGGGTGCCTGCGGCGCTCAT 573
Db 591 GGTCAACCGGAGGAGGCGCATCTTGGCGCTGCTCAGTGTCTGGTCTTGTCCACGCTCAT 650

QY 574 CGCCCTCGCGCGCTCTTGGCGGCGGAGGTGTGCGAGCGTTCGCGTCCAGCGGTG 633
Db 651 CTCATCGGCGCTCTCTTGGTGAAGGAGCGGCGCACCCACGAT---GACAAGGAGTG 707

QY 634 CCAGGTGAGCGGGAACCTCTATGCGCGCTTCTCCACCGCGCGGCTTCCACTGCC 693
Db 708 CGGGTCAACGAGAACCTCTTATGCGCCTCTTCTCTCTCTGCGCTCTTCTACATCCC 767

QY 694 GCTTGGCGTGTGCGTGTGTACCGGAAGATCTACGAGGCGGCGCAAGTGTTCGG 753
Db 768 TCTGGCGTGTCTTAGTCTAGTCTAGTCTAGTCTAGTCTAGTCTAGTCTAGTCTAG 827

QY 754 CCGCGCGGAGAGCTGTGCTGCGCTTTCGCGGCGGAGGTGTGCGAGCGTTCAGGTCAA 813
Db 828 GAACCTAGAGGAGGATCATGAAGGAGATGTCCAACTCCAAAGGAGTGAACCTG---A 883

```

```
QY 814 GGAAGCACCTGATGAGGCTGAAGTGTTCACGGCACATTGCAGAAGCAACGGTGTCTTT 873
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 884 GGATCCATTCCAGAACTTTTCAGGACACCCTTAGCAGTACCAAGGCCAAGGGCCACA 943
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 874 CCAGGTGAGCGGGGACTCCTGGCG-----GGAGCAGAAGGAGGCGAGCAGCCCA 923
Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 944 ACCCCAGAGTTCCATAGCTGTCAAACTTTTAAAGTTCTCCAGGGAAAGAAAGCAGCTA 1003
Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 924 TGATGTGGGAATCTGATGGCGTGTGCTGTGCTGGATCCCTTCTTCCTGACGG 983
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1004 AGACGTTGGGCATTGTGGTGGTATGTCATCTTGTGCTGGCTACCCCTTCTTCATCGTC 1063
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 984 AACTCATCAGCCCACTCTGTGCTGCAGCTGCCCC---CATCTGAAAAGCATATTC 1040
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1064 TACCGCTTGGCTCTTGTCTCCACCCTGAAGCCCCCGACGCCGTGTTCAAGGTGGTGT 1123
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1041 TGTGGCTTGGCTACTCCAATTTCTTCTCAACCCCTGATTTACACAGCTTTTAAACAAGA 1100
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1124 TCTGGCTGGCTACTTCAACAGCTGCTCAACCCCATCATCTACCCATGCTCCAGCAAGG 1183
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1101 ACTACAACAAATGCTTCAAGAGCTCTT 1128
Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 1184 AGTTCAAGCGCGTTTCGTGGCATCCT 1211
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
```

Search completed: September 7, 2004, 17:15:10
Job time : 114 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 7, 2004, 14:14:05 ; Search time 494 Seconds
(without alignments)
9906.729 Million cell updates/sec

Title: US-09-976-782-15

Perfect score: 1152

Sequence: 1 cgcacgagcgccgtagcc.....aagcagagatgaacacacagg 1152

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 3373863 seqs, 2124099041 residues

Total number of hits satisfying chosen parameters: 6747726

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

N Geneseq_29Jan04.*

1: Geneseqn1980s.*

2: Geneseqn1990s.*

3: Geneseqn2000s.*

4: Geneseqn2001as.*

5: Geneseqn2001bs.*

6: Geneseqn2002s.*

7: Geneseqn2003as.*

8: Geneseqn2003bs.*

9: Geneseqn2003cs.*

10: Geneseqn2004s.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1152	100.0	1152	6	ABN86917 Human NOV
2	1068	92.7	1150	4	ABA01984 Human NOV
3	1064.8	92.4	1150	4	ABA01985 Human NOV
4	1064	92.4	1458	6	AAD29681 Human G-p
5	1060	92.0	1119	6	AAD24963 Human G-p
6	1016	88.2	1155	6	ABK71923 Human cDN
7	756.2	65.6	3086	9	ADC86156 Human GPC
8	745	64.7	977	6	AAS98053 Human DNA
9	732.8	63.6	2036	2	AAQ70264 Murine se
10	732.8	63.6	2061	7	AAQ48747 Mouse 5-H
11	729.8	63.4	2226	7	AAQ72270 Rat MR22
12	729.8	63.4	2240	9	ADB52419 Primary r
13	721.2	62.6	1090	6	AAS98145 Human DNA
14	437.6	43.2	1074	6	ABK50434 Human 5-h
15	497.6	43.2	1074	7	ABZ42828 Human 5-H
16	497.6	43.2	1159	9	AAD58496 Human 5-h
17	496	43.1	2700	6	ABK34848 Human cDN
18	492.8	42.8	1074	5	ABi97963 Non-endog
19	484.4	42.0	1073	2	AAQ56309 Human SHI
20	435.2	37.8	1112	6	ABQ43624 Oligonucle
21	435.2	37.8	1112	6	ABQ43625 Oligonucle
22	424.6	36.9	1686	6	ABK13748 Mouse 5-H
23	421.4	36.6	1686	2	AAQ56308 Murine 5H

24	416	36.1	2155	2	AAQ72269 Rat REC17
25	390.2	33.9	1112	6	ABQ43623 Oligonucle
26	390.2	33.9	1112	6	ABQ43622 Oligonucle
27	378.2	32.8	636	4	AAH42175 Nucleotide
28	334.4	29.0	336	2	AAQ70265 Human ser
29	327.8	28.5	419	6	AAS98116 Human DNA
30	320	27.8	796	3	AAQ71992 Single nu
31	319.6	27.7	796	3	AAQ71995 Single nu
32	319.6	27.7	796	3	AAQ71989 Single nu
33	318.4	27.6	2543	6	ABK50432 Human 5-h
34	311.4	27.0	354	7	ABZ81690 Human LP3
35	297	25.8	297	4	AAH42174 Nucleotide
36	259.4	22.5	273	7	ABZ81688 Polynucle
37	211.4	18.4	225	7	ABZ81689 Polynucle
38	203.6	17.7	1929	6	ABK50433 Human 5-h
39	142.4	12.4	200	7	AAH48749 Targettin
40	125.4	10.9	1140	2	AAT88392 Corn barn
41	106	9.2	2826	7	ABZ42625 Human alp
42	104.4	9.1	1382	2	AAQ64890 Human der
43	104.4	9.1	1382	6	ABZ35643 Human gen
44	104.4	9.1	1382	7	ACA56658 Human sig
45	102.8	8.9	1371	4	AAI99933 Human alp

ALIGNMENTS

RESULT 1

ABN86917

ID ABN86917 standard; cDNA; 1152 BP.

AC ABN86917;

XX

DT 29-JUL-2002 (first entry)

XX

DE Human NOV5 encoding cDNA sequence SEQ ID NO:15.

XX

XX

KW Human; NOVX; cytostatic; antiarteriosclerotic; cardiovascular; lymphoma;

KW antidiabetic; immunosuppressive; neuroprotective; gene therapy; cancer;

KW cardiomyopathy; atherosclerosis; cell signal processing; diabetes; AIDS;

KW metabolic pathway modulation; neoplastic; neurological disorder; asthma;

KW adenocarcinoma; prostate cancer; uterus cancer; immune response;

KW Crohn's disease; multiple sclerosis; Graft versus host disease;

KW chromosome 2; gene; ss.

XX Homo sapiens.

OS

FH Key

FT CDS

FT Location/Qualifiers

5..1144

/*tag=

/product= "NOV5"

/note= "Serotonin receptor-like protein"

WO200230974-A2.

18-APR-2002.

XX

PF 12-OCT-2001; 2001WO-US031922.

XX

12-OCT-2000; 2000US-0240113P.

16-OCT-2000; 2000US-0240625P.

PR 16-OCT-2000; 2000US-0240637P.

PR 16-OCT-2000; 2000US-0240648P.

PR 16-OCT-2000; 2000US-0240662P.

PR 16-OCT-2000; 2000US-0240703P.

PR 16-OCT-2000; 2000US-0240732P.

PR 16-OCT-2000; 2000US-0241130P.

PR 18-JAN-2001; 2001US-0262455P.

XX

(CURA-) CURAGEN CORP.

PA (MILL/) MILLET I.

XX

Db 406 CTGTGCTGCCCCGCGCTCGGAAAGTGCGGCATCGCCCTGGCGCCGACGGGGCC 465
QY 485 ATCACAGGACCTGAGACAGACGCTGGCACCGGAGCGCGCTCGTTGCTCATGATC 544
Db 466 ATCACAGGACCTGAGACAGACGCTGGCACCGGAGCGCGCTCGTTGCTCATGATC 525
QY 545 GCGCTCGCCCGGGTGGCGCTCATCGCCCTCGCGCGCTGCTCTTTGGCGGGGCG 604
Db 526 GCGCTCGCCCGGGTGGCGCTCATCGCCCTCGCGCGCTGCTCTTTGGCGGGGCG 585
QY 605 GAGGTGTGCGACGCTGGCTCCAGCGCTGCCAGGTGAGCCGGAAACCTCTATGCCGCC 664
Db 586 GAGGTGTGCGACGCTGGCTCCAGCGCTGCCAGGTGAGCCGGAAACCTCTATGCCGCC 645
QY 665 TTCTCCACCGCGGCGCTTCCACCTCGCGCTGGCGTGGTGGTGGTGGTGGTGGTGG 724
Db 646 TTCTCCACCGCGGCGCTTCCACCTCGCGCTGGCGTGGTGGTGGTGGTGGTGGTGG 705
QY 725 ATCTACGAGGGGCGCAAGTTTCTGTTTGGCGCGCGCGAGAGCTGTGCTGCGTTGCCG 784
Db 706 ATCTACGAGGGGCGCAAGTTTCTGTTTGGCGCGCGCGAGAGCTGTGCTGCGTTGCCG 765
QY 785 GCCACATGAGGTGAGTCCAAAGTAAAGGAAGACCTGATGAGGCTGAAGTGGTGTTC 844
Db 766 GCCACATGAGGTG-----AAGGAAGACCTGATGAGGCTGAAGTGGTGTTC 813
QY 845 ACGGCACATGCAAGACAGCGTGTCTTCCAGGTGAGCGGGGACTCCTGGCGGGAGCAG 904
Db 814 ACGGCACATGCAAGACAGCGTGTCTTCCAGGTGAGCGGGGACTCCTGGCGGGAGCAG 873
QY 905 AAGGAGAGGCGAGCGCATGATGTTGGGAATTTCTGATTGGCGTGTGTTGCTGTCTGG 964
Db 874 AAGGAGAGGCGAGCGCATGATGTTGGGAATTTCTGATTGGCGTGTGTTGCTGTCTGG 933
QY 965 ATCCGCTTCTTCTGACGAACTCATPACGCCACTGTGCTGCGAGCTGCCCCCATC 1024
Db 934 ATCCGCTTCTTCTGACGAACTCATPACGCCACTGTGCTGCGAGCTGCCCCCATC 993
QY 1025 TGGAAAGCATATTTCTGGCTTGGCTACTCCAAATTTCTTCAACCCCTGATTTAC 1084
Db 994 TGGAAAGCATATTTCTGGCTTGGCTACTCCAAATTTCTTCAACCCCTGATTTAC 1053
QY 1085 ACAGCTTTTAAAGAACTACAACTGCTTCAAGAGCTCTTTACTAAGCAGAGATGA 1144
Db 1054 ACAGCTTTTAAAGAACTACAACTGCTTCAAGAGCTCTTTACTAAGCAGAGATGA 1113
QY 1145 ACACAGGG 1152
Db 1114 ACACAGGG 1121

RESULT 5
AAD24963
ID AAD24963 standard; cDNA; 1119 BP.
XX
AC AAD24963;
XX
DT 12-MAR-2002 (first entry)
XX
DE Human G-protein coupled receptor-8 (GREC-8) cDNA.
XX
KW Human; G-protein coupled receptor-8; GREC-8; therapy; cancer; stroke;
KW cell proliferative disorder; neurological; epilepsy; Parkinson's disease;
KW Alzheimer's disease; inflammation; thyroiditis; haemolytic anaemia; AIDS;
KW Acquired Immune Deficiency Syndrome; dementia; nontropic; cholelithiasis;
KW multiple sclerosis; atherosclerosis; angina pectoris; gastroenteritis;
KW diabetes; ulcer; viral infection; immunosuppressive; es.
XX
OS Homo sapiens.
XX
FH Key
XX Location/Qualifiers
FT 1. .1119
CDS

FT /*tag= a
FT /product= "Human GREC-8 protein"
FT 1. .195
FT sig_peptide /*tag= b
FT 196. .1116
FT mat_peptide /*tag= c
FT /*product= "Mature GREC-8 protein"
XX WO200198351-A2.
PN
XX
PD 27-DEC-2001.
XX
XX 15-JUN-2001; 2001WO-US019275.
XX
PR 16-JUN-2000; 2000US-0212483P.
PR 22-JUN-2000; 2000US-0213954P.
PR 29-JUN-2000; 2000US-0215209P.
PR 07-JUL-2000; 2000US-0216595P.
PR 14-JUL-2000; 2000US-0218936P.
PR 19-JUL-2000; 2000US-0219154P.
PR 21-JUL-2000; 2000US-0220141P.
XX (INCY-) INCYTE GENOMICS INC.
PI Lal P, Baughn MR, Hafalia AJA, Nguyen DB, Gandhi AR, Kallick DA;
PI Griffin JA, Yue H, Khan FA, Patterson C, Lu DAM, Tribouley CM;
PI Lu Y, Wallia NK, Graul R, Yao MG, Yang J, Ramkumar J, Au-Young J;
PI Elliott VS, Hernandez R, Walsh RT, Borowsky ML, Thornton M, He A;
XX WPI; 2002-075627/10.
DR P-PSDB; AAE15638.
XX
XX Isolated human G-protein coupled receptor polypeptides and the use of
XX these sequences in the diagnosis, treatment and prevention of diseases
XX and in the assessment of exogenous compounds on the expression of the
XX receptors.
PS Claim 11; Page 136-137; 143pp; English.
XX
XX The invention relates to isolated human G-protein coupled receptor
XX (GCREC) polypeptides and their biologically active fragments. GCREC and
XX protein is useful in treating a disease or condition associated with an
XX increase or decrease in expression of functional GCREC. The GCREC's are
XX useful in the diagnosis, treatment and prevention of cell proliferative
XX disorders (cancer, leukaemia, melanoma); neurological disorders (stroke,
XX epilepsy, Parkinson's disease, dementia; Alzheimer's disease); autoimmune
XX inflammatory disorder (thyroiditis, haemolytic anaemia, AIDS, multiple
XX sclerosis); cardiovascular disorder (atherosclerosis, angina pectoris),
XX gastrointestinal disorder (ulcer, cholelithiasis, gastroenteritis),
XX metabolic disorders (diabetes); viral infections (herpes virus) and in
XX the assessment of the effects of exogenous compounds on the expression of
XX the nucleic acid and amino acid sequences. The present sequence is human
XX GREC-8 cDNA
XX
SQ Sequence 1119 BP; 171 A; 376 C; 346 G; 226 T; 0 U; 0 Other;
Query Match 92.0%; Score 1060; DB 6; Length 1119;
Best Local Similarity 97.7%; Pred. No. 6.1e-220;
Matches 1114; Conservative 0; Mismatches 5; Indels 21; Gaps 3;
QY 5 ATGAGAGCGCTAGCCTTTTCAGTGGCCACCGCGCGGTTCCTTGGCCTGGACCCGAG 64
Db 1 ATGAGAGCGCTAGCCTTTTCAGTGGCCACCGCGCGGTTCCTTGGCCTGGACCCGAG 60
QY 65 ACACAGCGGACCGCGGACCCAGCCAGCGAGGATCTCGTTCCGACCCGAGCGGC 124
Db 61 ACACAGCG-----CGGACCCAGCCAGCGAGGATCTCGTTCCGACCCGAGCGGC 114
QY 125 GCCGTCTCTGCCGGCGGAGGGCGCCCTTCTCTGTCTTTCACGCTCTCTGTGTGAGCGCTG 184
Db 115 GCCGTCTCTGCCGGCGGAGGGCGCCCTTCTCTGTCTTTCACGCTCTCTGTGTGAGCGCTG 174
QY 185 CTAGTGTCTCTGATCGTCCCACTTTCCTGTGTGAACTCTGTGTTCGGTCCACCATCCCG 244

Db 175 CTAGTGTGCTGATCGCGCCACTTTCTCTGTGGAACCTGCTGTGGTTCGGTCAACCATCCG 234
Qy 245 CGGGTCCGTGCTTCCACCGGCTGCGGATACCTTGTGGCTCGAGCGGCTGTCCGAC 304
Db 235 CGGGTCCGTGCTTCCACCGGCTGCGGATACCTTGTGGCTCGAGCGGCTGTCCGAC 294
Qy 305 GAACCTAGTGCAGCGCTGCGGATGCCACCGAGCTGCGAGTGTGAGCTGTGACCGGCGGA 364
Db 295 GAACCTAGTGCAGCGCTGCGGATGCCACCGAGCTGCGAGTGTGAGCTGTGACCGGCGGA 354
Qy 365 CGTGGCTGTGCGCGCGAGCTGTGCCACGTGTGGATCTCTTCGACGCCGAGCCGTGT 424
Db 355 CGTGGCTGTGCGCGCGAGCTGTGCCACGTGTGGATCTCTTCGACG-----TG 405
Qy 425 CTGTGCTGCGCGCGCGCTCGGAAAGTGTGGCGGCATCGCCCTGGGCGCGAGCGGGCC 484
Db 406 CTGTGCTGCGCGCGCGCTCGGAAAGTGTGGCGGCATCGCCCTGGGCGCGAGCGGGCC 465
Qy 485 ATCACCGGACCTGCGAGCACAGCTGCGCACCGCGAGCGCGCTCGTTGCTCATGATC 544
Db 466 ATCACCGGACCTGCGAGCACAGCTGCGCACCGCGAGCGCGCTCGTTGCTCATGATC 525
Qy 545 GCGCTCGCGCGGCTGCGGCTCATCGCCCTGCGCGCTGCTGCTTTTGGCGCGGCG 604
Db 526 GCGCTCACCGGCTGCGGCTCATCGCCCTGCGCGCTGCTGCTTTTGGCGCGGCG 585
Qy 605 GAGTGTGCGAGCTGCGGCTCCAGCGCTCCAGTGTGAGCGGGAACCTCTATGCGGCC 664
Db 586 GAGTGTGCGAGCTGCGGCTCCAGCGCTCCAGTGTGAGCGGGAACCTCTATGCGGCC 645
Qy 665 TTCTCCACCGCGCGCTTCCACTGCGGCTTGGCGGTGCGGTTTGTCTACCGGAAG 724
Db 646 TTCTCCACCGCGCGCTTCCACTGCGGCTTGGCGGTGCGGTTTGTCTACCGGAAG 705
Qy 725 ATCTACGAGCGCGCAAGTTTCTGTTTGGCGCGCGCGGAGAGCTGTGCTCGGTTGCG 784
Db 706 ATCTACGAGCGCGCAAGTTTCTGTTTGGCGCGCGCGGAGAGCTGTGCTCGGTTGCG 765
Qy 785 GCGACCATGAGGTGAGTCCAGGTAAAGAGACCTGATGAGGCTGAAGTGGTGTTC 844
Db 766 GCGACCATGAGGTG-----AAGGTAAAGAGACCTGATGAGGCTGAAGTGGTGTTC 819
Qy 845 ACGGCACATTCAAGCAACGGTGTCTTCCAGGTGAGCGGGACTCTCGCGGAGAGCAG 904
Db 820 ACGGCACATTCAAGCAACGGTGTCTTCCAGGTGAGCGGGACTCTCGCGGAGAGCAG 879
Qy 905 AAGGAGAGCGGAGCAGCATGATGTTGGGAATTCTGATTGGCGTGTGTTGCTGTGTTG 964
Db 880 AAGGAGAGCGGAGCAGCATGATGTTGGGAATTCTGATTGGCGTGTGTTGCTGTGTTG 939
Qy 965 ATCCCTCTTCTGAGCGGAACCTATCAGCCCACTCTGTGCTCGAGCTGCGCCCATC 1024
Db 940 ATCCCTCTTCTGAGCGGAACCTATCAGCCCACTCTGTGCTCGAGCTGCGCCCATC 999
Qy 1025 TGGAAAGACATATTCTGTGCTTGGCTTGGCTTCCAACTCTTCTTCAACCCCTGATTAC 1084
Db 1000 TGGAAAGACATATTCTGTGCTTGGCTTGGCTTCCAACTCTTCTTCAACCCCTGATTAC 1059
Qy 1085 ACAGCTTTTAAAGAACTACAAAGTCCCTTCAAGAGCTCTTTTAAAGAGATGA 1144
Db 1060 ACAGCTTTTAAAGAACTACAAAGTCCCTTCAAGAGCTCTTTTAAAGAGATGA 1119

RESULT 6
ABK71923
ID ABK71923 standard; cDNA; 1155 BP.
XX
AC ABK71923;
XX
DT 30-JUL-2002 (first entry)
XX
DE Human cDNA encoding hydroxytryptamine receptor-like protein NOV8.

Human; ss; gene; NOVX; developmental disorder; endocrine disorder;
vascular disorder; infectious disease; anorexia; cancer; stroke;
neurodegenerative disorder; Alzheimer's disease; acute brain injury;
central nervous system disorder; depression; lung disorder;
reproductive disorder; tissue disorder; thrombocytopaenia; migraine;
angiogenesis; asthma; X-linked severe combined immunodeficiency;
inflammation; autoimmune disorder; immune disorder; blood disorder;
haematopoietic disorder; gastrointestinal disease; respiratory disorder;
hepatitis; fertility; hypertension; arteriosclerosis; ischaemia;
rheumatoid arthritis; Grave's disease; wound healing.

Homo sapiens.
WO200224733-A2.
28-MAR-2002.
17-SEP-2001; 2001WO-US029115.
15-SEP-2000; 2000US-0232675P.
15-SEP-2000; 2000US-0232676P.
15-SEP-2000; 2000US-0232679P.
18-SEP-2000; 2000US-023382P.
18-SEP-2000; 2000US-0233402P.
19-SEP-2000; 2000US-0233521P.
19-SEP-2000; 2000US-0233522P.
19-SEP-2000; 2000US-0233801P.
20-SEP-2000; 2000US-0233960P.
06-OCT-2000; 2000US-0238398P.
13-OCT-2000; 2000US-0240284P.
13-OCT-2000; 2000US-0240498P.
11-JAN-2001; 2001US-0260973P.
26-JAN-2001; 2001US-0264274P.
09-MAR-2001; 2001US-0274862P.

(CURA-) CURAGEN CORP.
Mishra VS, Syptek KA, Taupier RJ, Vernet CAM, Colman SD;
Gorman L, Tchernev VT, Malvankar UM, Shenoy S, Tchernev VT;
Padigaru M, Paturajan M, Burgess CE, Smithson G, Mallet I;
Peyman JA, Stone D, Gunther E, Ellerman K;
WPI; 2002-383182/41.
P-PSDB; ABG60235.

New cytoplasmic, nuclear, membrane bound and secreted NOVX polypeptides,
useful for treating cancers and tumors, lung disorders, hematopoietic
disorders, autoimmune diseases and immune disorders.

Claim 9; Page 59-60; 210pp; English.

The invention relates to an isolated NOVX polypeptide selected from
NOV1a, NOV1b, NOV1c, NOV2a, NOV2b, NOV3a, NOV3b, NOV4a, NOV4b,
NOV5a, NOV5b or NOV6-NOV9 polypeptides, their mature form or variant.
Also included are a nucleic acid encoding a NOVX protein or variant; a
vector comprising the nucleic acid; a cell comprising the vector; an anti-
NOVX antibody; and identifying agents that modulate the expression or
activity of NOVX. NOVX, the nucleic acid, antibody and modulators are
useful in the diagnosis, treatment or prevention of developmental
disorders, endocrine disorders, vascular disorders, infectious disease,
anorexia, cancer, neurodegenerative disorders (e.g. Alzheimer's disease,
Parkinson's disease, Huntington's disease, multiple sclerosis and
amyotrophic lateral sclerosis), acute brain injury (e.g. stroke, head
injury and cerebral palsy), central nervous system disorders (e.g.
depression, epilepsy and schizophrenia), lung disorders, reproductive
disorders, disorders affecting carbohydrate metabolism (e.g.
galactosaemia and hereditary fructose intolerance), tissue disorders
(e.g. Wiskott-aldrich syndrome, thrombocytopaenia, night blindness and
Pick's disease), disorders linked to abnormal angiogenesis, asthma,
azoospermia, learning disabilities, facial dysmorphism, autoimmune
encephalomyelitis, X-linked severe combined immunodeficiency, seizures,
migraines, inflammation, autoimmune disorders, disorders affecting sleep,

Qy	1	CGCATGAGGCGCGCTAGCTTTCACTGAGGCAACGCGCGGCTTGCCCTTGCCCTGGGACC	60
Db	1798	CGCATGAGGCGCGCTAGCTTTCACTGAGGCAACGCGCGGCTTGCCCTTGCCCTGGGACC	1857
Qy	61	CGAGACAGCAGCAGC-GGACCCGGGACCCAGCCGAGAGGATACTCGTTTCGACCCCGA	119
Db	1858	CGAGACAGCAGCAGGACCCGGGACCCAGCCGAGAGGATACTCGTTTCGACCCCGA	1917
Qy	120	GCGCGCCGCTCTGCGCGGCGGAGGGCGCCCTTCTCTGCTTTCAAGGNCCTCGTGCTGA	179
Db	1918	GCGCGCCGCTCTGCGCGGCGGAGGGCGCCCTTCTCTGCTTTCAAGGNCCTCGTGCTGA	1977
Qy	180	CGCTGTAGTGCTGCTGATCGCTGCCACTTTCCGTGTGAAACCTGCTGTTTCCCGTCAACA	239
Db	1978	CGCTGTAGTGCTGCTGATCGCTGCCACTTTCCGTGTGAAACCTGCTGTTTCCCGTCAACA	2037
Qy	240	TCCCGCGGCTCCGTCCTTCCACCGCGTGCAGATAACTTTGTGTGGCTCGACGCGCGTCT	299
Db	2038	TCCCGCGGCTCCGTCCTTCCACCGCGTGCAGATAACTTTGTGTGGCTCGACGCGCGTCT	2097
Qy	300	CGGACGAACTAGTGGCAGCGTGGCGATGCCACGAGCCTGGCGAGTGAGCTGTGACCG	359
Db	2098	CGGACGAACTAGTGGCAGCGTGGCGATGCCACGAGCCTGGCGAGTGAGCTGTGACCG	2157
Qy	360	GGCAGCTCGCTGCTGGCGGAGCCTGTGCCACAGTGTGATCTCTCTCGAGCCCGGAG	419
Db	2158	GGCAGCTCGCTGCTGGCGGAGCCTGTGCCACAGTGTGATCTCTCTCGAGCCCGGAG	2217
Qy	420	CTGTG-----CTGTGTGCCCGCGCGCTCGGGAAAGCTG	454
Db	2218	CTGTGCGCAGGTGTGGATCTCTTCCACGGCTGTGTCTGCCCGCGCGCTCGGGAAAGCTG	2277
Qy	455	GGGACCATCGCTCGGCGCGAGCGGGCCATCACAGCACCTTCGACGACACAGCTCGCG	514
Db	2278	GGGACCATCGCTCGGCGCGAGCGGGCCATCACAGCACCTTCGACGACACAGCTCGCG	2337
Qy	515	ACCGCAGCCGCGCTCTGTGCTCATATGCGCTCGCCCGGCTGCGGTGCGGCTCATC	574
Db	2338	ACCGCAGCCGCGCTCTGTGCTCATATGCGCTCGCCCGGCTGCGGTGCGGCTCATC	2397
Qy	575	GCCCTCGGCGCTGCTTTTGGCGGGGCGAGGTGTGCAGCCTCGGCTCCAGCGCTGC	634
Db	2398	GCCCTCGGCGCTGCTTTTGGCGGGGCGAGGTGTGCAGCCTCGGCTCCAGCGCTGC	2457
Qy	635	CAGGTGAGCCGGGAAACCTCTATGCGCGCTTCTCCACCCCGCGCGGCTTCCAGCTCGCG	694
Db	2458	CAGGTGAGCCGGGAAACCTCTATGCGCGCTTCTCCACCCCGCGCGGCTTCCAGCTCGCG	2517
Qy	695	CTTGGCGTGTGCGTTTGTCTACGGGAAGATCTACGAGGCGGCGAAGTTTCGTTTCGGC	754
Db	2518	CTTGGCGTGTGCGTTTGTCTACGGGAAGATCTACGAGGCGGCGAAGTTTCGTTTCGGC	2577
Qy	755	CGCGCCGAGAGCTGTGCTGCCGTGTGCGGCCACCATGCAGGTGAGTCCAAAGTAAAG	814
Db	2578	CGCGCCGAGAGCTGTGCTGCCGTGTGCGGCCACCATGCAGGTGAGTCCAAAGTAAAG	2637
Qy	815	GAAGCACT	823
Db	2638	AACGTTGCT	2646

RESULTS

RESULT 8
AAS98053
ID AAS98053 standard: DNA; 977 BP.

XX
AC AAS98053;

XX
DT 12-MAR-2002 (first entry)

XX Human DNA for potential G protein-coupled receptor #11. DE

XX
KW Human; G protein-coupled receptor; GPCR; ds; GAL4; galanin receptor;
KW Alzheimer's disease; amvotrophic lateral sclerosis; asthma;

atherosclerosis; basal cell carcinoma; breast carcinoma; cardiomyopathy; chondrosarcoma; chronic obstructive pulmonary disease; Crohn's disease; depression; epilepsy; macular degeneration; lymphoma; melanoma; multiple sclerosis; osteoarthritis; osteoporosis; Parkinson's disease; psoriasis; rheumatoid arthritis; schizophrenia; ulcerative colitis; tuberculosis; cognition disorder; memory disorder; anorexia; hormonal release disorder; cardiovascular activity disorder; pain perception disorder; obesity; diabetes; hypertension; diabetes; hyperlipidaemia; stroke; gene therapy.

Homo sapiens.

WO200185791-A1.

15-NOV-2001.

17-MAY-2001: 2001WO-US015332.

11-MAY-2000: 2000US-0203217P.

II-MAY-2000; 2000US-0205945P:
18-MAY-2000: 2000US-0205945P:

(LIFE-) LIFESPAN BIOSCIENCES INC.

Brown JP, Miller M, Burner G, Fabre-Suver C, Pritchard D;

WPT: 2002-066595/09.

Novel G protein-coupled receptor polypeptides including galanin receptor polypeptides useful for identifying modulators that are useful for treating Alzheimer's disease, psoriasis, melanoma, multiple sclerosis, stroke.

Disclosure; Page 119-120; 144pp; English.

The invention relates to an isolated polypeptide encoded by a nucleic acid molecule that is at least 80% identical to the G protein-coupled receptor (GPCR) polynucleotides included in the specification. Also included are probes based on the GPCR sequences (including antisense probes), a host cell comprising an expression vector comprising the GPCR sequence, antibodies raised against the polypeptides, and methods of identifying modulators of the polypeptides. The polypeptides are useful for identifying modulator compounds which function as modulators, activators, repressors, agonists or antagonists of the novel GPCR polypeptides including the GAL4 polypeptide. The antibodies and nucleic acid probes as described above can be used to detect the presence of the polypeptides and nucleic acids and are used to diagnose a variety of diseases or disorders in which GPCRs are involved e.g., Alzheimer's disease, amyotrophic lateral sclerosis, asthma, atherosclerosis, basal cell carcinoma, breast carcinoma, cardiomyopathy, chondrosarcoma, chronic obstructive pulmonary disease, Crohn's disease, depression, epilepsy, macular degeneration, lymphoma, melanoma, multiple sclerosis, osteoarthritis, osteoporosis, Parkinson's disease, psoriasis, rheumatoid arthritis, schizophrenia, ulcerative colitis, tuberculosis and many other diseases listed in the specification. The probes and antibodies are also useful for diagnosing cognition and memory disorders, anorexia, hormonal release disorders, cardiovascular activity disorders, pain perception disorders, obesity, diabetes, Alzheimer's disease. Preferably, compounds that decrease or increase the expression of galanin receptor (GAL4) can be used to treat obesity, diabetes, hyperlipidaemia and stroke. The GPCR nucleic acid is useful for treating the above mentioned disorders by gene therapy techniques. The present sequence is a novel GPCR polynucleotide of the invention.

Sequence 977 BP: 121 A; 356 C; 324 G; 176 T; 0 U; 0 Other; ;

Query Match 64.7%: Score 745: DB 6: Length 977;

Query Match 84.7%; SCORE 745; DB 6;
Best Local Similarity 94.6%; Pred. No. 9.6e-152;

BEST LOCAL SIMILARITY	94.0%;	ACC. NO. 3.00
MATCHES 803;	CONSERVATIVE 0;	MISMATCHES 20;
INDELS 26;	GAPS 26;	GAPS 26;

Qy 1 CGCCATGGAGGCGCTAGCCTTTCACTGGCCACCGCCGGCGTTGCCCTTGCCTCTGGGACC 60

Dh 85 CGCCGTGAGGCGCGCTAGCCTTTCACTGGCCACCGCCGGCGTTGCCCTTGCCTCTGGGACC 144

QY 61 CGAGACACGACG - GAGACCGGGACCCCAAGCCCGAGAGGGATCTCGGTTCGACCCCGA 119
 Db 145 CGAGACACGACGACGAGACCCGGACCCCAAGCCCGAGAGGGATCTCGGTTCGACCCCGA 204
 QY 120 GGGGCGCGCTCTGCGCGGCGGAGGCGCGCCCTTCTCTCTCTTACGGTCTCTGGTGA 179
 Db 205 GGGGCGCGCTCTGCGCGGCGGAGGCGCGCCCTTCTCTCTCTTACGGTCTCTGGTGA 264
 QY 180 CGCTGCTAGTGTCTGCTGCTGCTGCCACTTCTCTGGAACCTGCTGGTTCGGTCAACA 239
 Db 265 CGCTGCTAGTGTCTGCTGCTGCGCCACTTCTCTGGAACCTGCTGGTTCGGTCAACA 324
 QY 240 TCCCGGGTCCGCTGCTTCCACCGGCTGCCGATAACTTGGTGGCTCGACGGCGTCT 299
 Db 325 TCCCGGGTCCGCTGCTTCCACCGGCTGCCGATAACTTGGTGGCTCGACGGCGTCT 384
 QY 300 CGGACGAACTAGTGGCAGCGCTGGCGATGCCACCGAGCTGCGAGTGAAGTGTGACCG 359
 Db 385 CGGACGAACTAGTGGCAGCGCTGGCGATGCCACCGAGCTGCGAGTGAAGTGTGACCG 444
 QY 360 GGGCAGCTCGGCTGCTGGGCGGAGCTGTGCGACGTGGATCTCTTCGACGCGCGAG 419
 Db 445 GGGCAGCTCGGCTGTGGGCGGAGCTGTGCGACGTGGATCTCTTCGACGCGCGAG 504
 QY 420 CCGT-----CTGTGCTGCCCCCGCGCTCGGGAACTG 454
 Db 505 CCGTGTGCCACGTGTGATCTCCTTCCACGGCTGTGCTGCCCGCGCGCTCGGGAACTG 564
 QY 455 GGGGCACTCGCTGGGCGGAGCGGCGCATCACAGGCACTGCGACACAGCTGGCG 514
 Db 565 GGGGCACTCGCTGGGCGGAGCGGCGCATCACAGGCACTGCGACACAGCTGGCG 624
 QY 515 ACCCGCAGCGCGCTCGTGTCTCATGATGCGGCTCGCCCGGCTGCGGCGCTCATC 574
 Db 625 ACCTCGACCGCGCTCGTGTCTCATGATGCGGCTCACCCGGTGGCGTGGCGCTCATC 684
 QY 575 GCGCTCGCGCGCTGCTCTTTGGCGGCGGAGGTGTGCGACGCTCGGCTCCAGCGCTGC 634
 Db 685 GCGCTCGCGCGCTGCTCTTTGGCGGCGGAGGTGTGCGACGCTCGGCTCCAGCGCTGC 744
 QY 635 CAGGTGAGCGGAACTCTATGCGGCTTCTCACCGCGGCGCTTCACTTCACTGCGG 694
 Db 745 CAGGTGAGCGGAACTCTATGCGGCTTCTCACCGCGGCGCTTCACTTCACTGCGG 804
 QY 695 CTGCGTGTGCGCTTGTCTACCGGAAGATCTACGAGGCGGCAAGTTTCTGTTTCGC 754
 Db 805 CTGCGTGTGCGCTTGTCTACCGGAAGATCTACGAGGCGGCAAGTTTCTGTTTCGC 864
 QY 755 CGCGCGCGAGCTGTGCTGCGTTGCGGCGCACCATGCGAGTGAAGTGAAGTGAAG 814
 Db 865 CGACCGCGAGAGTGTGCTGCGGTTGCGGCGCACCATGCGAGTGAAGTGAAGTGAAG 924
 QY 815 GAAGCACT 823
 Db 925 AACGTTGCT 933

RESULT 9

AAQ70264

ID AAQ70264 standard; cDNA; 2036 BP.

XX AC AAQ70264;

XX DT 25-MAR-2003 (revised)

DT 27-MAR-1995 (first entry)

XX DE Murine serotonergic receptor 5HT5b cDNA.

XX KW Serotonergic receptor 5HT5b; 5-hydroxytryptamine; neuromodulator;
 KW mouse serotonin receptor; neurological; cardiovascular; psychiatric;
 KW disorder; ds.
 XX OS Mus musculus.

XX Key Location/Qualifiers
 EH 312..1424
 FT /*tag= a
 FT /product= "5HT5b"
 FT /note= "serotonin receptor"
 XX W09418319-Al.
 XX 18-AUG-1994.
 XX 07-FEB-1994; 94WO-FR000136.
 XX 09-FEB-1993; 93FR-00001392.
 XX (INM) INST NAT SANTE & RECH MEDICALE.
 XX Amlaiky N, Boschert U, Grailhe R, Hen R, Matthes H, Plassat J;
 PI WPI; 1994-279740/34.
 XX P-PSDB; AAE57066.
 XX New serotonergic receptor 5HT5b and related nucleic acid and
 PT recombinant cells - useful for treatment and diagnosis of e.g.
 PT neurological or cardiovascular disease, also for identifying serotonin
 PT agonists and antagonists.
 XX Claim 7; Page 14-16; 27pp; French.
 XX Oligonucleotides corresponding to conserved regions of known serotonin
 CC receptors were used in PCR of mouse brain RNA in presence of reverse
 CC transcriptase and the products were sequenced. One product with homology
 CC to known receptors was labelled and used to probe a cDNA library in
 CC Lambda UnizAP. The 2036 bp sequence AAQ70264 was identified and deduced
 CC to encode a protein having 7 hydrophobic domains which represents a novel
 CC serotonin receptor which was designated 5HT5b. (Updated on 25-MAR-2003 to
 CC correct PN field.) (Updated on 25-MAR-2003 to correct PA field.)
 XX SQ Sequence 2036 BP; 425 A; 582 C; 542 G; 487 T; 0 U; 0 Other;
 Query Match 63.6%; Score 732.8; DB 2; Length 2036;
 Best Local Similarity 79.9%; Pred. No. 4.9e-149;
 Matches 911; Conservative 0; Mismatches 202; Indels 27; Gaps 3;
 QY 5 ATGAGGCGCTAGCTTTCTAGTGCCACCGCCGCTGCTTCCCTGGACCCCGAG 64
 Db 312 ATGGAAGTTTCTAACCTCTCAGGCCCACTCCCGCCCTTCTTCTCCGGGACTGAG 371
 QY 65 ACCAGCAGCGACCCGGGACCCCAAGCCCGAGAGGATACCTCGGTTGACACCCCGAGCGGC 124
 Db 372 AGCTGCAGTGAC-----AGCCCAAGTTCCGGCAGAGAGATGGGATCCACCCAGGTGGG 425
 QY 125 GCGTCTCTGCGGCGGAGGCGCGCTTCTGTCTTCAAGCTCTCGGTCTGCTGAGCGGTG 184
 Db 426 CTCATCTTSCCGCGCGGAGCGCGCTTCTGTCTTTCACCGCTGCTTGTGTGACTCTA 485
 QY 185 CTAGTGTCTGTATGCTGCTGCACCTTCTCTGTGGACCTGCTGTTCCGTCACCATCCCG 244
 Db 486 CTGTGTGCTGTATGCTGCTGCCACTTTCTATGGAATCTGTAGTCTTCTGGTACTATCTG 545
 QY 245 CGGTTCTGCTTCCACCGGCTGCCGATAACTTGGTGGCTCGACGCGCGTCTTCGGAC 304
 Db 546 CGCGTCCGCGCTTCCACCGGCTGCCGATAACTTGGTGGCTCGACGCGCGTCTTCGGAT 605
 QY 305 GAACTAGTGGCAGCGCTGGCGATGCCACCGAGCTGGCGAGTGAAGTGTTCACCGGGCGA 364
 Db 606 GTCTGTGTGGCGGTTCGTGTGATGCTCTGAGCTGTGTGAGCGAGTGTCTCGCTGGGCGA 665
 QY 365 CGTGGTGTGTGGCGGAGCTGTGCGACCTGTGGATCTCTCTCGACGCGCGAGCTGT 424
 Db 666 CGTTGGCAGCTAGGACGAGAGTCTGTGCGACCTGTGGATCTCTCTTCGAGC-----TG 716
 QY 425 CTGTGCTGCCCGCGGCTCGGGAAACGTGGCGGCGCATCGCCCTTGGGCGCGACGCGGCGC 484

Db 717 TTGTGTCACCGCCACCACTGGAACGTTGGGSCCATCGCCCTGGATCGTACTGGAAT 776
Qy 485 ATACACGGACCTGACAGCACACCTGCGACACCGGACCGCGCTCGTGTCTCATGATC 544
Db 777 ATACCGGCGACCTGACAGTACACCTGCGACCGGACCGGCTGCTTGTGGCTCATGATC 836
Qy 545 GCGTCTGCGCGGTCGCTGCGGCTCATCGCCCTCGCGGCTGCTTGTGGCGGGC 604
Db 837 GCGATCACCTGGGCACTGTGCGGCTCATGCTCTCGCCCGCTGCTTTTGGCTGGGC 896
Qy 605 GAGGTGTGCGACGCTCGGCTCCAGCGTGGCAGTGAGCGGGAACCTCTCATGCGGCC 664
Db 897 GAACCTATGATGCTCGGCTGCAGCGTGCAGGTGAGCCAGGAGCCCTCTCATGCTGC 956
Qy 665 TTCTCCACCGCGCGCTTCACACTGCGGCTTGGCGTGGTGGCTGCTTGTACCGAAG 724
Db 957 TTCTCCACCTGCGGAGCTTCTACCTGCTCTAGCGGTGGTGTCTTCTGCTACTGAAA 1016
Qy 725 ATCTACGAGGCGGCAAGTTTCTGTTTCGCGCGCGCGGAGAGCTGTGCGGTGCGG 784
Db 1017 ATATACAAAGCGGCAAGTTTCTGATTCTGCTGCGAGACGCGCGGCTGTACCGCTTCT 1076
Qy 785 GCCACCATGAGGTGAGTCTCAAGGTAAAGGAAGCACCTGATGAGGCTGAAGTGTTC 844
Db 1077 GCCACGAGC-----AGGCAAGGAAGCACCTCGGAGTCTGAGATGTTTC 1124
Qy 845 ACGGCAATGCAAGCAACAGGTGTCTTCCAGGTGAGCGGGGACTCTTGGCGGAGAG 904
Db 1125 ACAGCCGCTGCGGAGCAACAGTGACCTTCCAGACAAGCGGAGACTCTTGGCGGAGCAG 1184
Qy 905 AAGGAGGCGGAGCAGCATGATGTTGGGATTTCTGATGGCGTGTGCTGCTGCTGG 964
Db 1185 AAGGAGAAGCGGAGCAGCATGATGTTGGGATCTTGAATGGCGTGTGCTTGTGG 1244
Qy 965 ATCCCTCTTCTTCCGAGGAACTCATCAGCCCACTCTGTGCTCGAGCTGCCCGCCATC 1024
Db 1245 ATCCCTCTTCTTCCGAGGACTCATCAGCCCGCTCTGTGCTCGAGCTGCCCGCCATC 1304
Qy 1025 TGGAAAAGCATATTTCTGTGCTGGTACTTCCAAATTTCTTCAACCCCTGATTTAC 1084
Db 1305 TGGAAAAGCATATTTCTGTGCTGGATATTCCAATTCGTTCTTCAACCCCTGATTTAC 1364
Qy 1085 ACAGCTTTTACAGAACTACAACATGCTTCAAGACCTCTTACTAAGCAGATGA 1144
Db 1365 ACTGCTTTTAAAGATTACAACATGCTTCAAGACCTCTTACTAAGCAGATGA 1424

RESULT 10
AAD48747
XX AAD48747 standard; DNA; 2061 BP.
XX AC
XX AC
XX AC
DT 07-MAR-2003 (first entry)
XX Mouse 5-HT5B receptor gene.
XX
XX Mouse; 5-hydroxytryptophan; 5-HT5B receptor; depression; transgenic;
transgenic animal; phenotype; pain sensitivity; neurological disease;
gene therapy; pharmacological; neuropsychological disease; analgesic;
psychotic illness; receptor; gene; ds.
XX
OS Mus sp.
XX
XX Key Location/Qualifiers
FH misc_feature 72..391
FT /tag= a
FT /note= "Sequence flanking Neo insert in targeting
construct"
FT 312..1424
FT CDS /tag= b
FT /product= "Mouse 5-HT5B receptor protein"

FT misc_feature 392..464
FT /tag= c
FT /note= "Sequence deleted in targeting construct"
FT misc_feature 465..1008
FT /tag= d
FT /note= "Sequence flanking Neo insert in targeting
construct"
XX WO200279443-A2.
XX
XX 10-OCT-2002.
XX
XX 29-MAR-2002; 2002WO-US009853.
XX
XX 29-MAR-2001; 2001US-0280553P.
XX 21-DEC-2001; 2001US-0342472P.
XX 28-MAR-2002; 2002US-00109532.
XX
XX (DELT-) DELTAGEN INC.
XX
XX Allen KD;
XX
XX WPI; 2003-067450/06.
XX P-PSDB; AAE48747.
XX
XX Novel transgenic mouse useful for identifying a potential therapeutic
agent for the treatment of depression or pain, comprises disruption in a
5-hydroxytryptophan 5B receptor gene.
XX
XX Example 1; Fig 1; 60pp; English.
XX
XX The present invention relates to transgenic animals, compositions and
methods relating to the characterisation of gene function. The invention
also relates to transgenic mice comprising mutations in 5-hydroxy-
tryptophan (5-HT)5B receptor gene. The transgenic mice are useful for
identifying an agent that modulates the phenotype such as increased
depression or increased pain sensitivity. They are also useful for
identifying potential therapeutic agents for the treatment of pain or
depression and for evaluating a potential therapeutic agent capable of
affecting a condition associated with a mutation in a 5-HT5B receptor
gene. Transgenic animals of the invention are also useful for testing the
efficacy of proposed genetic and pharmacological therapies for human
diseases such as neurological, neuropsychological or psychotic illnesses.
XX The present sequence is mouse 5-HT5B receptor gene
XX
XX Sequence 2061 BP; 446 A; 582 C; 545 G; 488 T; 0 U; 0 Other;
Query Match 63.6%; Score 732.8; DB 7; Length 2061;
Best Local Similarity 79.9%; Pred. No. 4.9e-149;
Matches 911; Conservative 0; Mismatches 202; Indels 27; Gaps 3;
Qy 5 ATGGAGCGCGTAGCCTTTCAGTGGCCACCGCGCGTTCCTTCCCTCGGACCCGAG 64
Db 312 ATGGAAGTTTCTAACTCTCAGCGGCCACTCCCGGCTTGCCTTCTCCGGAGCTGAG 371
Qy 65 ACCAGCAGCGACCGCGGACCCCAAGCCCGAGAGGATCTCGTTCACCCCGAGCGGC 124
Db 372 AGCTGCAGTGAC-----AGCCCAAGTTCCCGCAGGAGCATGGGATCCACCCAGTGGG 425
Qy 125 GCCGTCTTCGCGGCGGAGCGCCGCTTCTGTCTTTCAGGCTCTGTGTGTGACGCTG 184
Db 426 CTCATCTTGGCGGCGGAGCGCGCTTCTCTGCTTTCACCGTGTGTGTGACTCTA 485
Qy 185 CTAGTGTGCTGATGCTGCGACCTTCTGTGGAACCTGCTGCTTCCGCTCACCATCCG 244
Db 486 CTGCTGTGCTGATGCTGCGACCTTCTTATGGAATCTGTAGTCTGTGTGACTATCTG 545
Qy 245 CGGCTCGCTGCTTCCACCGCGTCCGATAACTTGTGTGGCTCGACCGCGCTCTCGGAC 304
Db 546 CGGTCGCGGCTTCCACCGGCTGCCACATTAATTTGTGGCTCGACAGCGCTCTCGAT 605
Qy 305 GAACCTAGTGGAGCGCTGGCGATGCCACCGAGCCCTGGCGAGTGTGAGCTGTGACCGGGCGA 364

```
Db 606 GTCTGCTGGCGGTTCTGTGTGATGCTCTAGAGCTGTGTGAGCGAGTGTTCGCTGGCGGA 665
QY 365 CFTGCGCTCTGTGGCGGAGCCTGTGCCACGTTGTGATCTCCTTCGACGCGGAGCCTGT 424
Db 666 CGTGGCAGTAGCAGAGTCTGTGCCAGTGTGATCTCCTTCGAG-----TG 716
QY 425 CTGTGCTGCGCGCGCTCGGAACTGGGAACTGGGCGCATCGCCTGGCGCGGAGCGGCC 484
Db 717 TTGTGCTGACCGCCAGCATCTGAAAGTGGGCGCATCGCCTGGATCGTACTGGACT 776
QY 485 ATCACAGGCACTGTCAGCACAGCTGCGCACCGGACCGGAGCGGCGCTGTGTCTATGATC 544
Db 777 ATACAGCGCACCTGTCAGTACAGCTTGGCGACCGGAGCGGCTTCTCGCTCATGATC 836
QY 545 GCGCTGCGCGCGGTGCGCTCGGCGCTCATCGCCTCGCGCGCTGTCTTTGCGCGGCG 604
Db 837 GCGATCACCTGGCACTGTCCGGCTCATGTCTCGCGCGCTGTCTTTTGGCTGGGCG 896
QY 605 GAGGTGTGCGAGCTCGGCTCCAGCGCTCCAGGTCGCGAGTGGCGGAAACCTTCCTATGCGCC 664
Db 897 GAAGCCTATGATGCTCGGCTGAGCGTTCGCAAGTGGAGGAGCGCCTCCTATGCTGTC 956
QY 665 TTCTCCACCGCGGCGCTTCACCTGCGCTTGGCGTGTGGCGTGTCTTCTACCGAG 724
Db 957 TTCTCCACCTGCGGAGCCTTACCTGCTCTAGCGGTGTGTCTTCTGCTACTGAAA 1016
QY 725 ATCTAGAGGCGGCCAAAGTTTCGTTTCGGCGCGCGGAGAGCTGTGCTGCGTTCGCG 784
Db 1017 ATATACAAAGCCCAAGTTTCGATTCGTCGAGAGCGGCGGTGTACGCTTCT 1076
QY 785 GCACATGTCAGGTGAGTCCAAAGTAAAGAGACCTGTATGAGGCTGAAGTGGTTC 844
Db 1077 GCCACGCGC-----AGGCAAGAGAACCACTCCGAGTCTGAGATGGTGTTC 1124
QY 845 ACGGCACATTCGAAAGCAAGCTGTCCTTCAGGTCGAGCGGAGCTCTGCGGAGCAG 904
Db 1125 ACAGCCGTCGCGGAGACAGTACCTTCAGACAGCGGAGCTCTGCGGAGCAG 1184
QY 905 AAGGAGGCGGAGCAGCAGCATGATGTGGAAATCTGATTCGCTGTTTGTGCTGCTGG 964
Db 1185 AAGGAGAGCGGCGCAGCATGATGTCGGATCTTGATTCGCTGTTTGTGTTGG 1244
QY 965 ATCCCTTCTTCTGAGGAACTATCAGCCCACTGTGCTGCTGAGCCTGCCCCCATC 1024
Db 1245 ATCCCTTCTTCTGAGGAGCTATCAGCCCGCTCTGTGCTGAGCCTGCGCCCATC 1304
QY 1025 TGGAAAGCATATTTCTGTGGCTTGGTACTTCAATTCCTTCAACCCCTGATTTAC 1084
Db 1305 TGGAAAGCATATTTCTGTGGCTTGGATATTCATTCCTTCAACCCCTGATTTAC 1364
QY 1085 ACAGCTTTTAAAGAACTACAAATGCTTCAAGAGCCTTTTACTAAGCAGAGATGA 1144
Db 1365 ACTGCTTTTAAAGAACTTAAACATGCTTCAAGAGCCTTTTACTAAGCAGAGATA 1424
```

RESULT 11

```
AAQ72270
ID AAQ72270 standard; cDNA; 2226 BP.
XX
AC AAQ72270;
```

```
DT 25-MAR-2003 (revised)
DT 24-MAY-1995 (first entry)
```

```
XX
DE Rat MR22 serotonin receptor gene.
```

```
XX Serotonin; receptor; transmembrane; domain; kinase; phosphorylation;
KW sensory; motor; behaviour; central nervous system; CNS; superfamily;
KW G-protein; ligand-gated; ion channel; subfamily; human; rat; amplifi;
KW primer; PCR; amplification; brain; hypothalamus; indolamine; drug;
KW hypothalamus; therapeutic; neurological; pathology; dementia; insomnia;
KW Parkinson's disease; eating disorder; anxiety; migraine; headache; ss.
```

OS Rattus rattus.

```
XX Key Location/Qualifiers
CDS 303..1415
FT /tag= a
FT /product= "rat MR22 serotonin receptor"
```

```
XX WO9421670-A1.
```

```
XX 29-SEP-1994.
```

```
XX 15-MAR-1994; 94WO-US002839.
```

```
XX 15-MAR-1993; 93US-00031538.
```

```
XX (SCRI ) SCRIPPS RES INST.
```

```
XX Sutcliffe JG, Erlander MG, Lovenberg TW;
```

```
XX WPI; 1994-316932/39.
```

```
XX P-PSDB; AAR58686.
```

```
XX New serotonin receptors and corresp. DNA and antibodies - useful in
diagnosis and treatment of neurological processes and pathologies.
```

```
XX Claim 13; Page 137-138; 198pp; English.
```

```
XX The nucleotide sequence of the novel rat serotonin receptor MR22. The
gene encodes a protein of 370 amino acids. The protein contains 7
transmembrane domains (TMDs), a putative N-linked glycosylation site and
4 putative sites for kinase C phosphorylation. Serotonin regulates a wide
range of sensory, motor and behavioural functions in the central nervous
system (CNS). Serotonin receptors belong to at least two protein
superfamilies: G-protein-associated receptors containing 7 TMDs
(including 5-HT1a/b/c/d/e, 5-HT2) and ligand-gated ion channel receptors
with 4 TMDs (5-HT3). The serotonin receptors presented in the patent
represent prototypes that fall into three new serotonin subfamily
classifications: 5-HT1e-like (rat MR77 - AAQ72271 and human MR77 -
AAQ72272), 5-HT5 (subdivided into 5-HT5alpha, rat REC17 - AAQ72269 and 5-
HT5beta, rat MR22 - AAQ72270) and 5-HT6 (rat RBC20 - AAQ72273). The genes
for the receptors were cloned by amplification based on conserved amino
acid sequences found in the TMDs. Degenerate primers were constructed to
these sequences and used in two rounds of nested PCR amplification on a
rat brain hypothalamic cDNA template. In the second round of
amplification, degenerate primers corresponding to conserved residues
found only in indolamine-binding receptors were used to amplify only
those receptors. The resultant fragments were cloned into pBluescript
vectors. The plasmid fragments were used to screen a rat hypothalamus
cDNA library to obtain full length clones. The nucleic acids molecules
and their corresponding proteins may be used in methods for determining
ligand binding activity, detecting and altering expression of serotonin
receptors, drug screening and therapeutic treatments involving human
serotonin receptors. Serotonin receptors are also known to participate in
neurological processes. Pathologies including dementia, Parkinson's
disease, eating disorders, pathological anxiety, migraine, headaches,
insomnia and other conditions. (Updated on 25-MAR-2003 to correct PN
field.)
```

```
XX Sequence 2226 BP; 488 A; 641 C; 592 G; 505 T; 0 U; 0 Other;
```

```
XX Query Match 63.4%; Score 729.8; DB 2; Length 2226;
XX Best Local Similarity 79.6%; Pred. No. 2.2e-148;
XX Matches 911; Conservative 0; Mismatches 207; Indels 27; Gaps 3;
```

```
QY 3 CCATGGAGCGCTAGCCTTTTCAGTGGCCAGCGCGGTGGCTTCCCTGGGACCGG 62
Db 301 CAATGGAGTCTCTAACTCTCAGCGCCACCCCTGGCATTCCTTCTCGGACCGG 360
QY 63 AGACCAGCAGCGGACCCCGGAGCCCAAGCCGAGAGGATATCTCGTTTCGACCCCGAGG 122
Db 361 AGAGCTGAGTGAC-----AGCCCAAGTTCGCGAGAGCATGCCATCCACCCAGGG 414
QY 123 GCGCGTCTCTGCGGCGCGGAGCGCGCTTCTTGTCTTTCACGTCCTGTGGTGAACG 182
```

415	Db	GGCTCATCTTTGTCGGCGCGGAGCGGCCCTTCTGCTCTCACCGTACTCGTGTAATC	474
183	Qy	TGCTAGTGCTGATCGCTGCCACTTTCGTGTGGAACTGCTGTTCCGGTCACCATCC	242
475	Db	TACTGGTGTGCTGATCGCTGCCACTTTCATTATGGAATCTGCTAGTCTCTGGTGAATACC	534
243	Qy	CGGGGTCCGTGCTTCCACCGGTGCGCATAACTTGGTGGCTCGACGGCCGTCTCGG	302
535	Db	TGCGGTCGCGCCCTTCCACCGGTGCGCATAACTTGGTGGCTCGACGGCCGTCTCGG	594
303	Qy	ACGAACTAGTGGACGCGCTGGCGATGCCACGGAGCTGGCGAGTGACTGTGCACGGGC	362
595	Db	ACGTCCTGTGTGGCGCTCTGGTATGCCACTGAGCCCTGGTGAAGAGTGTTCGGCTGGC	654
363	Qy	GAGTCGCTGTGCGGCGGAGCTGTGCCACGTGTGATCTCTCTCGACGGCCGAGCCCT	422
655	Db	GACGTTGGAGCTGGGACGAGTCTGTGCCACGTGTGATCTCTCTCGACG-----	705
423	Qy	GTCTGTCTGCCCGCGGCTTCGGGAACTGTGGCGCATGTGCCCTGGGCGGACGGG	482
706	Db	TGTTGTGCTGCACAGCCAGCATCTGGAACGTGGCGGCATCGCCCTGGATCTGGA	765
483	Qy	CCATCACGCGCACTTCAGACACGCTGGCACCGCAGCCGCGCCTCGTGTCTATGA	542
766	Db	CTATCACGCGCACTTCAGTACAGCTGGGACCGCGCGCGCTTCAGCACTTATGA	825
543	Qy	TCGCGCTCGCCCGGCTGCGTTCGCGCTCATCGCCCTCGCGCCCTGCTTTGGCGCGG	602
826	Db	TCGCGATCACTGGGCACTGTCCGCTCATCGGCTCGCCCGCTGCTTTGGCTGGG	885
603	Qy	GCGAGGTGTGCGACGCTCGGCTCAGCGCTGCCAGGTGAGCGCGGAACCTCTTATCGG	662
886	Db	GCGAAGCCTACGATGCTCGGCTGAGCGTTCGCAAGTGGAGCAGGAGCTTCGTAACCGG	945
663	Qy	CTTCTCCACCGCGGCGCTTCACCTGCGCTTGGCGTGGTGGCTTGTCTACCGGA	722
946	Db	TCCTTCTCCACCTGCGGAGCTTTCAGTGTGCTCTGGCGGTGGTGTCTTCTGCTACTGGA	1005
723	Qy	AGATCTACGAGCGCGCAAGTTTCGTTTCGGCGCGCGCGGAGAGCTGTGCTGCCGTTCG	782
1006	Db	AGATATACAAAGCCGCAAGTTTCGATTCGCGCGCGAGACGGCGGCGGTAGTGCCTGCG	1065
783	Qy	CGGCCACCATGAGGTGAGTTCGAAGGTAAAGGAAGCACTGTATGAGCTGAAGTGTGT	842
1066	Db	CCGCCACCAAGC-----AGGCAAGGAAGCACTTCAGAGTCTGAGACGGTAT	1113
843	Qy	TCAGGCACATTGCAAGCAAGCTGCTCTTCAGGTGAGCGGGAGCTTCGTGGCGGAGC	902
1114	Db	TCACGCGCGTGTGAGAGCGACAGTGGCTTCGAGACGAGTGGAGACTCTTGGCGGAGC	1173
903	Qy	AGAAGGAGGAGCGAGCAGCCATGATGTGGGAATTCGATTGGGCTGTTGTGTGTGTCT	962
1174	Db	AGAAGGAGGAGCGAGCGCCATGATGTGGGATCTTGATCGGTGTGTTGTGTGTGTCT	1233
963	Qy	GGATCCCTTCTTCTGACGGAATCATCAGGCCACTGTGTGCTGAGACGCTGCCCCCA	1022
1234	Db	GGATCCCTTCTTCTGACGGAAGTCTGTGAGCGCGCTCTGAGCGCTGCAACCCA	1293
1023	Qy	TCTGGAAAGCATATTTCTGTGGCTTGGCTACTCAATTTCTTCAACCCCTGATTT	1082
1294	Db	TCTGGAAAGCATATTTCTGTGGCTTGGCTATTTCAATTCGTTTCTCAATCCCTTAATCT	1353
1083	Qy	ACACAGCTTTTAAACAGAACTACAACAATGCTTCAAGAGCCTCTTTAATAAGCAGAGAT	1142
1354	Db	ACACGCGCTTTTAAAGAACTACAACAATGCTTCAAGAGCCTCTTTAATAAGCAGAGAT	1413
1143	Qy	GAACA 1147	
1414	Db	AAGCA 1418	

ADB52419
ID ADB52419 standard; DNA; 2240 BP.
XX
XX
AC ADB52419;
XX
DT 04-DEC-2003 (first entry)
XX
XX
DE Primary rat hepatocyte toxicity modelling related gene SEQ ID NO:2961.
XX
XX
KW toxic effect; gene expression profile; hepatotoxicity; diagnostic marker;
KW toxicity marker; toxicity progression; drug screening;
KW primary rat hepatocyte toxicity modelling; gene; ds.
XX
XX
OS Rattus norvegicus.
XX
XX
FN WO2003065993-A2.
XX
XX
PD 14-AUG-2003.
XX
XX
PF 04-FEB-2003; 2003WO-US003482.
XX
XX
PR 04-FEB-2002; 2002US-0353171P.
PR 13-MAR-2002; 2002US-0363534P.
PR 08-APR-2002; 2002US-0370248P.
PR 10-APR-2002; 2002US-0371134P.
PR 10-APR-2002; 2002US-0371135P.
PR 10-APR-2002; 2002US-0371150P.
PR 11-APR-2002; 2002US-0371413P.
PR 19-APR-2002; 2002US-0373601P.
PR 19-APR-2002; 2002US-0373602P.
PR 22-APR-2002; 2002US-0374139P.
PR 08-MAY-2002; 2002US-0378370P.
PR 09-MAY-2002; 2002US-0378652P.
PR 09-MAY-2002; 2002US-0378653P.
PR 09-MAY-2002; 2002US-0378665P.
PR 09-JUL-2002; 2002US-0394230P.
PR 09-JUL-2002; 2002US-0394253P.
PR 04-SEP-2002; 2002US-0407688P.
PR 28-JAN-2003; 2003US-0442900P.
XX
PA (GENE-) GENE LOGIC INC.
XX
XX
PI Mendrick D, Porter M, Johnson K, Higgs B, Castle A, Orr M;
PI Elashoff M;
XX
XX
DR WPI; 2003-731472/69.
XX
XX
PT Determining if a compound induces a toxic effect on a tissue or cell, for
PT identifying hepatotoxic compounds, comprises comparing a gene expression
PT profile of a tissue or cell sample to a database of Tox mean and non-Tox
PT mean values.
XX
XX
PS Claim 44; SEQ ID NO 2961; 874pp; English.
XX
XX
CC The present invention describes a method for determining whether a
CC compound induces a toxic effect on a tissue or cell. The method comprises
CC preparing a gene expression profile of a tissue or cell sample exposed to
CC the compound, and comparing the gene expression profile to a database
CC comprising data or information on the Tox mean and non-Tox mean value.
CC The method is useful for predicting or identifying at least one toxic
CC effect, particularly hepatotoxicity, of a test or unknown compound. The
CC genes listed in the specification are useful as diagnostic or toxicity
CC markers for the prediction or identification of the physiological state
CC of tissue or cell sample that has been exposed to a compound, or to
CC identify or predict the toxic effects of a compound or an agent. These
CC may also be used as markers for monitoring toxicity progression or for
CC drug screening. The present sequence represents a primary rat hepatocyte
CC toxicity modelling related gene sequence from the present invention.
XX
XX
SQ Sequence 2240 BP; 490 A; 646 C; 597 G; 507 T; 0 U; 0 Other;

Query Match 63.4%; Score 729.8; DB 9; Length 2240;
Best Local Similarity 79.6%; Pred No. 2.2e-148;

Matches 911; Conservative 0; Mismatches 207; Indels 27; Gaps 3;

```
QY 3 CCATGAGGCGGTAGCTTTTCAAGTGGCCACCGCGGCTTGCCTTGCCTCGGACCGG 62
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
301 CAATGGAAGTCTTAACTCTAGGCGCCACCCCTGGCATTTGCTTTCTCCGGGACCGG 360
QY ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
63 AGACACGACGCGACCCCGGACCCCAAGCCGAGAGGAGTACTCGGTTCCGACCCCGAGCG 122
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
361 AGAGCTGCAGTGAC-----AGCCCAAGTTCGGCAGAGCATGGGATCCACCCAGGGG 414
QY ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
123 GCGCGCTCTCGCGGCGGAGGCGCGCTTCTCTGTCTTTCAGGTCCTGGTGGTACGC 182
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
415 GGTCTATCTGTTCGGCGCGGAGCGCGCTTCTCGCTTCACTGCTACTGCTGTAATC 474
QY ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
183 TGCTAGTGTCTGTATCGCTGCGACATTTCTGTGGAACTGCTGGTTCGCTCAACATCC 242
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
475 TACTGGTGTGTATCGCTGCCACTTTCTATGGAATCTGTAGTCTGTGTACTATCC 534
QY ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
243 CGGGGTCGTGCTTCCACCGGTCGCGATTAACCTTGTGGCTCGAGCGCGCTCTCGG 302
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
535 TCGCGTTCGCGCTTCCACCGTGTGCCATAAATTTGGTAGCTCGACCGCGCTCTCGG 594
QY ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
303 AGCAACTAGTGGCAGCGCTGGGATGCCACCGAGCTGGGAGTGGCTGCACCGGCG 362
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
595 AGTCTGTGTGGCGCTCTGGTATGCCACTGAGCTGTGGAGCGAGTTGTGGCTGGCG 654
QY ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
363 GACGTGCGCTGTGGCGGAGCTGTGCCAGTGTGGATCTCTTCGACGCGCGGAGCCT 422
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
655 GAGCTTGGCAGCTGGCAGGAGTCTGTGCCAGTGTGGATCTCTTCGAGC----- 705
QY ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
423 GTCTGTGCTGCCCGCGCGCTTCGGGAACGTGGCGGCCATGCGCTTGGCGCGGAGCGG 482
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
706 TGTGTGTGTGCACAGCCAGCATCTGGAACGTGGCGGCCATCGCCCTGGATCGTACTGA 765
QY ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
483 CCATCACAGCGCACTGTGACACAGCTGTGGCACCGGACCGCGCTGTGCTCATCA 542
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
766 CTATACGGCGCACCTGTGAGTACAGCTGCGCACCGCGCGCGCTTCAGCACTATGA 825
QY ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
543 TCGCGCTCGCGCGGTGCGTGGCGCTCATCGCCCTCGCGCGCTGTCTTTGGCGGG 602
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
826 TCGCGATCACCTGGCACTGTCCGCCCTCATCGCTCGCCCGCTGTCTTTGGCTGG 885
QY ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
603 GCGAGTGTGCAAGCTGTGGTTCAGCGCTGCAGGTGAGCGCGGAACCTCTCTATGCG 662
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
886 GCGAAGCCCTACGATGCTCGCTGCGAGCTTGCAGGTGAGCGAGCGCTTGTACTGCG 945
QY ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
663 CTTCTCTCCACCGCGCGCTTCCACCTGCGCTTGGCGTGTGCTTGTCTACCGGA 722
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
946 TCTTCTCACCTCGGAGCTTCTACGTGCTTGTGCGCTGTGCTTGTCTACTGA 1005
QY ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
723 AGATCTAGAGCGGCCAAGTTTCGTTTGGCGCGCGCGGAGAGCTGTGCTCGCGTTGC 782
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
1006 AGATATACAAAGCCGCAAGTTTCGATTCGGCGCGCAGAGCGCGGCTAGTGCCCTGC 1065
QY ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
783 CGGCCACCATGCAAGTGTCAAGTTAAGAGACCTGATGAGGCTGAAGTGGTGT 842
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
1066 CGGCCACCAAGC-----AGGCAAGAGAGACCTCAGAGTCTGAGCGGTAT 1113
QY ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
843 TCACGGCACATTCGAAAGCAAGTGTCTTCCAGTGTAGCGGGACTCTCGCGCGGAGC 902
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
1114 TCACCGCGTGTGAGAGGACAGTGGCTTCCACAGAGTGGAGACTCTCGCGCGGAGC 1173
QY ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
903 AGAAGGAGCGGAGCAGCCATGATGTTGGGAATCTGATTTGGCGTGTGTTGTGTGT 962
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
1174 AGAAGGAGAGCGAGCGCCATGATGTTGGGATCTTGATCGTGTGTTGTGTGTGT 1233
QY ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
963 GATCCCTCTTCTGACGGAATCATCAGCCACTGTGCTGAGCGCTGAGCGCGCCCA 1022
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
1234 GATCCCTCTTCTGACGAGCTGTGAGCGCGCTTGCAGCGCTGAGCGCTGAGCGCCCA 1293
QY ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
1023 TCTGAAAGCATATTTCTGTGGCTTGGCTACTCCAAATTTCTTCAACCCCTGATTT 1082
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
1294 TCTGAAAGCATATTTCTGTGGCTTGGCTATCAAAATCTGTTCAATCCCTTAATCT 1353
```

```
QY 1083 ACACAGCTTTTAAACAAGAACTACAACAATGCTTCAAGAGCCTCTTTACTAAGCAGAGAT 1142
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
1354 ACACGGCCTTTTAAAGAATACTACAACATGCTTCAAGAGCCTCTTTACTAAGCAGAGAT 1413
QY ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
1143 GAACA 1147
1414 AAGCA 1418
RESULT 13
AAS98145
ID AAS98145 standard; DNA; 1090 BP.
XX
AC AAS98145;
XX
DT 12-MAR-2002 (first entry)
XX
Human DNA for potential G protein-coupled receptor #102.
XX
Human; G protein-coupled receptor; GPCR; ds; GAL4; galanin receptor;
KW Alzheimer's disease; amyotrophic lateral sclerosis; asthma;
KW atherosclerosis; basal cell carcinoma; breast carcinoma; cardiomyopathy;
KW chondrosarcoma; chronic obstructive pulmonary disease; Crohn's disease;
KW depression; epilepsy; macular degeneration; lymphoma; melanoma;
KW multiple sclerosis; osteoarthritis; osteoporosis; Parkinson's disease;
KW psoriasis; rheumatoid arthritis; schizoprenia; ulcerative colitis;
KW tuberculosis; cognition disorder; memory disorder; anorexia;
KW hormonal release disorder; cardiovascular activity disorder;
KW pain perception disorder; obesity; diabetes; obesity; diabetes;
KW hyperlipidaemia; stroke; gene therapy.
XX
OS Homo sapiens.
XX
WO2001085791-A1.
XX
PD 15-NOV-2001.
XX
11-MAY-2001; 2001WO-US015332.
XX
11-MAY-2000; 2000US-0203217P.
PR 18-MAY-2000; 2000US-0205945P.
XX
(LIFE-) LIFESPAN BIOSCIENCES INC.
XX
Brown JP, Miller M, Burmer G, Fabre-Suver C, Pritchard D;
WPI; 2002-066595/09.
XX
Novel G protein-coupled receptor polypeptides including galanin receptor
PT polypeptides useful for identifying modulators that are useful for
PT treating Alzheimer's disease, psoriasis, melanoma, multiple sclerosis,
PT stroke.
XX
Claim 2; Page 103; 144pp; English.
XX
The invention relates to an isolated polypeptide encoded by a nucleic
CC acid molecule that is at least 80% identical to the G protein-coupled
CC (GPCR) polynucleotides included in the specification. Also included are
CC probes based on the GPCR sequences (including antisense probes), a host
CC cell comprising an expression vector comprising the GPCR sequence,
CC antibodies raised against the polypeptides, and methods of identifying
CC modulators of the polypeptides. The polypeptides are useful for
CC identifying modulator compounds which function as modulators, activators,
CC repressors, agonists or antagonists of the novel GPCR polypeptides
CC including the GAL4 polypeptide. The antibodies and nucleic acid probes as
CC described above can be used to detect the presence of the polypeptides
CC and nucleic acids and are used to diagnose a variety of diseases or
CC disorders in which GPCRs are involved e.g., Alzheimer's disease,
CC amyotrophic lateral sclerosis, asthma, atherosclerosis, basal cell
CC carcinoma, breast carcinoma, cardiomyopathy, chondrosarcoma, chronic
CC obstructive pulmonary disease, Crohn's disease, depression, epilepsy,
CC macular degeneration, lymphoma, melanoma, multiple sclerosis,
```


CC osteoarthritis, osteoporosis, Parkinson's disease, psoriasis, rheumatoid
CC arthritis, schizophrenia, ulcerative colitis, tuberculosis and many other
CC diseases listed in the specification. The probes and antibodies are also
CC useful for diagnosing cognition and memory disorders, anorexia, hormonal
CC release disorders, cardiovascular activity disorders, pain perception
CC disorders, obesity, diabetes, Alzheimer's disease. Preferably, compounds
CC that decrease or increase the expression of galanin receptor (GAL4) can
CC be used to treat obesity, diabetes, hyperlipidaemia and stroke. The GPCR
CC nucleic acid is useful for treating the above mentioned disorders by gene
CC therapy techniques. The present sequence is a novel GPCR polynucleotide
CC of the invention
CC
CC
XX
SQ Sequence 1090 BP: 157 A; 364 C; 372 G; 197 T; 0 U; 0 Other;

Query Match	62.6%;	Score 721.2;	DB 6;	Length 1090;
Best Local Similarity	95.2%;	Pred. No. 1.4e-146;		
Matches 775;	Conservative	0;	Mismatches 13;	Indels 26;
				Gaps 2

36	Qy	CCGGCGTTGGCCCTTGGCCCTGGGACCCGAGACACAGAGC-GGACCCGGGACCCCAAGCCGG	94
1	Db	CCGGCGTTGGCCCTTGGCCCTGGGACCCGAGACACAGAGCAGGACCCGGGACCCCAAGCCGG	60
95	Qy	AGAGGGATACTCGGTTCCAGACCCCGAGCGGCGCGTCTGCGGGCCGAGGGGCGGCCCTTC	154
61	Db	AGAGGGATACTCGGTTCCAGACCCCGAGCGGCGCGTCTGCGGGCCGAGGGGCGGCCCTTC	120
155	Qy	TCTGTCTTACGGTCTCGGTTGAGCGTGTAGTGTCTGATCGCTGCCACTTTCCTTG	214
121	Db	TCTGTCTTACGGTCTCGGTTGAGCGTGTAGTGTCTGATCGCTGCCACTTTCCTTG	180
215	Qy	TGGAACTGTGGTTCCGGTCAACCATCCGCGGGTCCGTTCCTTCCACGGGTGCGGCAT	274
181	Db	TGGAACTGTGGTTCCGGTCAACCATCCGCGGGTCCGTTCCTTCCACGGGTGCGGCAT	240
275	Qy	AACCTGTGGCCTTCGAGCGGCCGTCTCGGACGAACTAGTGGCAGCGCTGCCGATGCCACGG	334
241	Db	AACCTGTGGCCTTCGAGCGGCCGTCTCGGACGAACTAGTGGCAGCGCTGCCGATGCCACGG	300
335	Qy	AGCTCGCGAGTAGCTGTGCACCGGGCGACGTGGCTGTCTGGGCGGAGCGCTGTGCCAC	394
301	Db	AGCTCGCGAGTAGCTGTGCACCGGGCGACGTGGCTGTCTGGGCGGAGCGCTGTGCCAC	360
395	Qy	GTGTGGATCTCTTCGAGCGCGGAGCCGTG-----CTGTG	429
361	Db	GTGTGGATCTCTTCGAGCGCGGAGCCGTGTGCCAGTGTGGATCTCTTCCACGGCTGTG	420
430	Qy	CTGCCCGCGCGCCTTCGGGAACTGTGGCGGCCATCGCCCTGGGCGCGACGGGGCCATCAC	489
421	Db	CTGCCCGCGCGCCTTCGGGAACTGTGGCGGCCATCGCCCTGGGCGCGACGGGGCCATCAC	480
490	Qy	ACGGCACTTGAGACACAGCTTGGACCCGAGCCGCGCGCTCGTTGTCTCATGATCGCGT	549
481	Db	ACGGCACTTGAGACACAGCTTGGACCCGAGCCGCGCGCTCGTTGTCTCATGATCGCGT	540
550	Qy	CGCCCGGTGCGTGGCGCTCATCGCCTCGCGCGCTGCTCTTTGGCCGGCGGAGGT	609
541	Db	CGCCCGGTGCGTGGCGCTCATCGCCTCGCGCGCTGCTCTTTGGCCGGCGGAGGT	600
610	Qy	GTGCGACGCTCGGTTCACGCGCTGCAGGTGAGCCGGAAACCTCTATGCCGCTTTC	669
601	Db	GTGCGACGCTCGGTTCACGCGCTGCAGGTGAGCCGGAAACCTCTATGCCGCTTTC	660
670	Qy	CACCCGGCGCGCTTCACCTGCGCTTGGCGTGGTTCCTTCACCGGAGATCTA	729
661	Db	CACCCGGCGCGCTTCACCTGCGCTTGGCGTGGTTCCTTCACCGGAGATCTA	720
730	Qy	CGAGCGGCCCAAGTTTCGTTTTCGGCGCGCGCGGAGAGCTGTGCTGCCGCTTTC	789
721	Db	CGAGCGGCCCAAGTTTCGTTTTCGGCGCGCGCGGAGAGCTGTGCTGCCGCTTTC	780
790	Qy	CATGAGGTGAGGTCCTCAAGGTAAAGGACACT	823
781	Db	CATGAGGTGAGGTCCTCAAGGTAAAGGACACT	814

RESULT 14	
ABK50434	
ID	ABK50434 standard; DNA; 1074 BP.
XX	
AC	ABK50434;
XX	
30-JUL-2002	(first entry)
XX	
Human 5-hydroxytryptamine receptor 5A (HTR5A) DNA coding region.	
XX	
Human; 5-hydroxytryptamine receptor 5A; HTR5A; serotonin; gene; ds;	
KW	neuroprotective; neurological disease; depression; epilepsy;
KW	gene therapy; single nucleotide polymorphism; haplotype pair;
KW	chromosome 7q36.1.
XX	
Homo sapiens.	
XX	
Key	Location/Qualifiers
CDS	1..1074
FT	/*tag= a
FT	/product= "Human HTR5A"
FT	replace(12,T)
FT	/*tag= b
FT	/standard_name= "Single nucleotide polymorphism"
FT	replace(43,T)
FT	/*tag= c
FT	/standard_name= "Single nucleotide polymorphism"
FT	replace(300,C)
FT	/*tag= d
FT	/standard_name= "Single nucleotide polymorphism"
FT	replace(358,C)
FT	/*tag= e
FT	/standard_name= "Single nucleotide polymorphism"
FT	replace(459,G)
FT	/*tag= f
FT	/standard_name= "Single nucleotide polymorphism"
FT	replace(469,A)
FT	/*tag= g
FT	/standard_name= "Single nucleotide polymorphism"
FT	replace(789,T)
FT	/*tag= h
FT	/standard_name= "Single nucleotide polymorphism"
FT	replace(849,T)
FT	/*tag= i
FT	/standard_name= "Single nucleotide polymorphism"
XX	
WO200222887-A1.	
XX	
21-MAR-2002.	
XX	
17-SEP-2001; 2001WO-US0292210.	
XX	
15-SEP-2000; 2000US-0233051P.	
XX	
(GENA-) GENAISSANCE PHARM INC.	
XX	
Kazemi A, Koshy B, Sanchis A, Tirrell C;	
XX	
WPI; 2002-393978/42.	
XX	
P-PSDB; AAU79252.	
XX	
Novel genetic variants of 5-Hydroxytryptamine (Serotonin) Receptor 5A	
PT	isogenes, useful for improving efficiency and reliability in drug
PT	development for treating neurological diseases.
XX	
Claim 27; Fig 3; 134pp; English.	
XX	
The invention relates to single nucleotide polymorphisms in the gene	
CC	encoding human 5-hydroxytryptamine (serotonin) receptor 5A (HTR5A). A
CC	method for haplotyping the HTR5A gene in an individual comprises
CC	identifying the nucleotide at one or more polymorphic sites and

determining whether one of the copies of the gene is defined by one of the HTR5A haplotypes given in the specification or whether both copies are defined by a haplotype pair. This method is useful in genotyping, whereby all possible haplotype pairs can be assigned to specific genotypes. An association between a trait and a haplotype or haplotype pair of the HTR5A gene can be identified by comparing the frequency of the haplotype or haplotype pair in a population exhibiting the trait with the frequency of the haplotype or haplotype pair in a reference population, where a higher haplotype frequency in the trait population indicates the trait is associated with the haplotype or haplotype pair. HTR5A and its corresponding DNA are used for studying the expression and function of HTR5A, and in screening for candidate drugs to treat diseases related to HTR5A activity, such as neurological disorders, including depression and epilepsy. This sequence represents the DNA encoding the human HTR5A polypeptide.

Sequence 1074 BP: 182 A; 366 C; 284 G; 242 T; 0 U; 0 Other;

RESULT 15

ABZ42828

ID ABZ42828 standard: DNA: 1074 BP.

AC ABZ42828;

04-MAR-2003 (first entry)

Human 5-HT5A receptor nucleotide SEQ ID NO:445.

KW G protein-coupled receptor; GPCR, antigenic peptide; gene therapy;
KW G protein-coupled receptor; modulator; antibody; immune-related disease;
KW growth-related disease; cell regeneration-related disease; AIDS; cancer;
KW immunological-related cell proliferative disease; autoimmune disease;
KW Alzheimer's disease; atherosclerosis; infection; osteoarthritis; allergy;
KW osteoporosis; cardiomyopathy; inflammation; Crohn's disease; diabetes;
KW graft versus host disease; Parkinson's disease; multiple sclerosis; pain;
KW psoriasis; anxiety; depression; schizophrenia; dementia; memory loss;
KW mental retardation; epilepsy; asthma; tuberculosis; obesity; nausea;
KW hypertension; hypotension; renal disorder; rheumatoid arthritis; trauma;
KW ulcer; gene; ds.

XX Homo sapiens.

PN WO200261087-A2.

08-AUG-2002.

AA
PF
19-DEC-2001: 2001WO-US050107.

PR 19-DEC-2000; 2000US-0257144P.

PA (LIFE-) LIFESPAN BIOSCIENCES INC.

PI Burmer GC, Roush CL, Brown JP;

DR WPI; 2003-046718/04.

DR P-PSDB; ABP81980.

New isolated antigenic peptides e.g., for G protein-coupled receptors (GPCR), useful for diagnosing and designing drugs for treating conditions in which GPCRs are involved, e.g. AIDS, Alzheimer's disease, cancer or autoimmune diseases.

PS Disclosure; Fig 1; 523pp; English.

CC The present invention describes antigenic peptides (I) comprising: (a)

CC any one of 1601 sequences (see ABP82019 to ABP83619) of 12-24 amino

CC acids. Also described: (1) an assay for the detection of a particular G

CC protein-coupled receptor (GPCR) or a candidate polypeptide in a sample; CC and (2) an isolated antibody having high specificity and high affinity for

CC and (2) an isolated antibody having high specificity and high affinity or

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 7, 2004, 15:11:05 ; Search time 2862 seconds
(without alignments)
12019.997 Million cell updates/sec

Title: US-09-976-782-15

Perfect score: 1152

Sequence: 1 cgcacatgagccgtagcc.....aagcagatgacacacaggg 1152

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 27513289 segs, 14931090276 residues

Total number of hits satisfying chosen parameters: 55026578

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

EST:*

1: em_estba:*

2: em_esthum:*

3: em_estin:*

4: em_estmu:*

5: em_estcov:*

6: em_estpl:*

7: em_estro:*

8: em_htc:*

9: gb_est1:*

10: gb_est2:*

11: gb_htc:*

12: gb_est3:*

13: gb_est4:*

14: gb_est5:*

15: em_estfun:*

16: em_estom:*

17: em_gss_hum:*

18: em_gss_inv:*

19: em_gss_pln:*

20: em_gss_vrt:*

21: em_gss_fun:*

22: em_gss_mam:*

23: em_gss_mus:*

24: em_gss_pro:*

25: em_gss_rod:*

26: em_gss_phg:*

27: em_gss_vrl:*

28: gb_gss1:*

29: gb_gss2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	496	43.1	1074	29	AY402561	AY402561 Homo sapi
2	495	43.0	1074	29	AY402562	AY402562 Pan trogl
c 3	477.6	41.5	640	28	BZ601879	BZ601879 WHAAAS5TF
4	426.2	37.0	1074	29	AY402563	AY402563 Mus muscu

5	284.2	24.7	1099	12	BM805859	BM805859 AGENCOURT
6	280.4	24.3	906	12	BI489745	BI489745 603032245
7	265.4	23.0	539	13	BY280849	BY280849 BX280849
8	255.6	22.2	1027	29	CNS05MTV	AL344524 Tetraodon
c 9	255.2	22.2	893	12	BI603546	BI603546 603244448
c 10	255.2	22.2	2683	11	BC031272	BC031272 Homo sapi
c 11	247.6	21.5	583	12	BI538349	BI538349 428889 MA
c 12	239.6	20.8	472	12	BG715343	BG715343 602677361
c 13	200.4	17.4	929	12	BI411238	BI411238 602964578
c 14	198.4	17.2	660	10	BB628489	BB628489 BB628489
c 15	173.8	15.1	655	12	BM861710	BM861710 fy49e02.x
c 16	149.4	13.0	316	14	CD600820	CD600820 RK125A4B0
c 17	148.2	12.9	744	29	CE490868	CE490868 tigr-gss-
c 18	144.2	12.5	646	12	BM861408	BM861408 fy49e02.y
c 19	140.8	12.2	730	12	BI820920	BI820920 603033562
c 20	120.2	10.4	538	9	AL922333	AL922333 AL922333
c 21	112.6	9.8	874	14	CB565883	CB565883 AGENCOURT
c 22	107	9.3	697	10	BB653529	BB653529 BB653529
c 23	96.6	8.4	1213	29	AY402238	AY402238 Pan trogl
c 24	93.6	8.1	988	29	CNS018EH	AL109235 Drosophil
c 25	93.6	8.1	1269	29	AY402237	AY402237 Homo sapi
c 26	93	8.1	503	14	CB784054	CB784054 AMGNNUC:U
c 27	93	8.1	745	12	BG827135	BG827135 602751117
c 28	93	8.1	780	12	BI195997	BI195997 602756310
c 29	93	8.1	914	13	BQ879192	BQ879192 AGENCOURT
c 30	93	8.1	929	13	BQ719305	BQ719305 AGENCOURT
c 31	93	8.1	1332	29	AY418851	AY418851 Homo sapi
c 32	91.4	7.9	519	28	QO897350	QO897350 HS_5444_A
c 33	91.4	7.9	1335	29	AY418853	AY418853 Mus muscu
c 34	91.4	7.9	1201	9	AL530418	AL530418 AL530418
c 35	90	7.8	1058	12	BM926157	BM926157 AGENCOURT
c 36	88.8	7.7	1221	29	AY409288	AY409288 Pan trogl
c 37	88.8	7.7	1323	29	AY417222	AY417222 Homo sapi
c 38	86.6	7.5	1201	9	AL544609	AL544609 AL544609
c 39	86.4	7.5	1221	29	AY409287	AY409287 Homo sapi
c 40	84.4	7.3	1764	11	BC018330	BC018330 Mus muscu
c 41	84.4	7.3	2611	11	AK004891	AK004891 Mus muscu
c 42	84.4	7.3	3050	11	AK043877	AK043877 Mus muscu
c 43	83.8	7.3	1072	29	AY398880	AY398880 Homo sapi
c 44	83	7.2	1401	29	AY399426	AY399426 Homo sapi
c 45	82.4	7.2	1125	29	AY418220	AY418220 Mus muscu

ALIGNMENTS

RESULT 1	AY402561	1074 bp	DNA	linear	GSS 15-DEC-2003
LOCUS	Homo sapiens HTR5A gene, VIRTUAL TRANSCRIPT, partial sequence,				
DEFINITION	genomic survey sequence.				
ACCESSION	AY402561				
VERSION	AY402561.1	GI:39758547			
KEYWORDS	GSS.				
SOURCE	Homo sapiens (human)				
ORGANISM	Homo sapiens				
REFERENCE	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;				
AUTHORS	Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.				
	1 (bases 1 to 1074)				
	Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejarawal,A.,				
	Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B.,				
	Ferriera,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J.,				
	Adams,M.D. and Cargill,M.				
TITLE	Infering nonneutral evolution from human-chimp-mouse orthologous				
JOURNAL	gene trios				
PUBMED	Science 302 (5652), 1960-1963 (2003)				
REFERENCE	14671302				
AUTHORS	2 (bases 1 to 1074)				
	Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejarawal,A.,				
	Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B.,				
	Ferriera,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J.,				
	Adams,M.D. and Cargill,M.				
TITLE	Direct Submission				

Qy	953	GTGCTGCTGGATCCCTTTCTTCTGA	CGGAACATCATCAGCCACCTCTGTGCTGCAGC	1012
Db	883	GTGCTCTGTCTGGATCCCTTTCTTCT	CACCGAGCTCATCAGTCCCTCTGCTCTGTGAC	942
Qy	1013	CTGCCCCCATCTGGAAAGCATATTCT	GTGCTGGCTTGGCTACTCCAAATCTTTCTTCAAC	1072
Db	943	ATCCCGCCCATCTGGAAAGCATCTTC	CTGTGGCTTGGCTACTCCAACTCTTCTTTAAC	1002
Qy	1073	CCCTGATTACACAGCTTTTAAACAAG	ACTACAAGTCGCTTCAAGACCTCTTTTACT	1132
Db	1003	CCCTGATCTATACGGCTTTTCAACA	AGAACTACAAGCGCTTCAAGAACTCTTTTCT	1062
Qy	1133	AAGCAGAGATGA	1144	
Db	1063	AGGCAACACTGA	1074	
RESULT 2				
AY402562		1074 bp	DNA	linear
LOCUS				GSS 15-DEC-2003
DEFINITION	AY402562	Pan troglodytes HTR5A gene, VIRTUAL TRANSCRIPT, partial sequence,		
ACCESSION	AY402562	genomic survey sequence.		
VERSION	AY402562			
KEYWORDS	AY402562.1	GI:39758548		
SOURCE		GSS.		
ORGANISM		Pan troglodytes (chimpanzee)		
REFERENCE		Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Pan.		
AUTHORS		Clark,A.G., Gланowski,S., Nielson,R., Thomas,P., Kejariwal,A., Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B., Ferreira,S., Wang,G., Zheng,X.H., White,T.J., Sainsky,J.J., Adams,M.D. and Cargill,M.		
TITLE		Inferring nonneutral evolution from human-chimp mouse orthologous gene trios		
JOURNAL		Science 302 (5652), 1960-1963 (2003)		
PUBMED		14671302		
REFERENCE		2 (bases 1 to 1074)		
AUTHORS		Clark,A.G., Gланowski,S., Nielson,R., Thomas,P., Kejariwal,A., Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B., Ferreira,S., Wang,G., Zheng,X.H., White,T.J., Sainsky,J.J., Adams,M.D. and Cargill,M.		
TITLE		Direct Submission		
JOURNAL		Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive, Rockville, MD 20850, USA		
COMMENT		This sequence was made by sequencing genomic exons and ordering them based on alignment.		
FEATURES		Location/Qualifiers		
source		1..1074		
		/organism="Pan troglodytes"		
		/mol_type="genomic DNA"		
		/db_xref="taxon:9598"		
gene		<1..>1074		
		/gene="HTR5A"		
		/locus_tag="HCM1265"		
ORIGIN				
Query Match		43.0%	Score 495;	DB 29; Length 1074;
Best Local Similarity		70.4%;	Pred. No. 3.3e-80;	
Matches		727; Conservative 0;	Mismatches 281;	Indels 24; Gaps 4;
Qy	113	ACCCGAGCGCGCGCTCTGTCGCGGGCGGAGGCGCGCTTCCTGTCTCAGCGTCTG	172	
Db	67	AGCCTCGCAAGACGACCTCGCGCCCGAGCTCGCGCTCTTCGGGTCTTCGAGTGCTT	126	
Qy	173	GTGGTAGCGTGTAGTGTCTGATCGTGCACATTTCTCTGTGGAACTGCTGTTC	232	
Db	127	ATTCTACCTTGTCTGGGCTTCTGGTGGCGGCGACGTTCCCTTGGAACTGCTGTGGTCTG	186	
Qy	233	GTCAACATCCCGGGGTCCGTGCTTCCACCGGCTGCGCATAACTTGGTGGCTCGACG	292	

```

Db 187 GCGACCATCTCCGTGTACGACCTTCCACCGGTGCCCCCAACACCTGGTGCATCCATG 246
Qy 293 GCGCTCGGAGCACTAGTGGCAGCGTGGCGATGCCACCGAGCTGGCAGTGCAGCTG 352
Db 247 GCGATCTGGATGCTCTGTGGCCGCGTGGTATGCGCGCTGAGCTGGTGACAGCTG 306
Qy 353 TCGACCGGCGACGTGGTGTGGCCGCGAGCTGTGCGACGTGTGGATCTCTTCGAC 412
Db 307 T---CGGCGCGCTGGCAGCTAGTGGAGGCTGTGCCAGCTTGGATCTCGTGGAC 363
Qy 413 GCGGAGCTGTCTGTGCTGCCCGCGCGCTGGGAGCTGGCGCATGCGCTTGGGC 472
Db 364 G-----TGCTTTGTGTCAGCGCCAGCATCTGGAACGTGACGCCATAGCCCTGGAC 414
Qy 473 GCGGAGCGGCGCATCACAGCGCACTGCGAGCACAGCTGCGCACCGCGAGCGCGCTCG 532
Db 415 CGCTACTGTGCATCAGCGCGCCACATGAATACAGCTCCGACCCCGAAGTGGCTGCC 474
Qy 533 TTGCTCATGATCGCGCTCGCCCGGGTGGCGTGGCGGCTCATCGCCCTGCGCGCTGCTC 592
Db 475 AAGGTCATGATCGCGCTCACCTGGGCACTCTCCGCTGTCTCTCTGCGCCGCTGCTT 534
Qy 593 TTGCGCGGCGGAGTGGAGCGCTCGCTCGCTCGAGCGCTGCGAGTGGAGCGCGGAAACC 652
Db 535 TTGCTGGGAGAGACGCTACTCTGAGGCGAGCGAGGAGTGCAGGTAAAGCGCGAGGCT 594
Qy 653 TCCTATGCGCGCTCTCCACCGCGCGCTTCCACCTGCGCTTGGGTGGTGGTGGTGGT 712
Db 595 TCCTAGCGCGTCTCTCACCGTGGCGCTTCTACCTGCGCTGCTGGTGGTGGTGGTGGT 654
Qy 713 GTCTACCGGAAGATCTACGAGCGCGCAAGTTTCGTTTCGCGCGCGCGGAGAGCTGTG 772
Db 655 GTGTAAGTGAAGATCTACAGGCTGCCAAGTTCGCGGTGCGGCTCCAGGAGACCAATAGC 714
Qy 773 CTGCGGTCGCGGCGACCATGACGTGAGGTCCAAAGTTAAAGGAAGCACTGATGAGCT 832
Db 715 GT-----CTCACCCCATATCCGAAGCTGTGGAGTGAAGGANTCTGCCAAACAGGCC 765
Qy 833 GAAGTGGTGTTCACGGCACATTTGAAAGCAAGGTGTCTTCCAGGTGAGCGGGAGCTCC 892
Db 766 CAGATGGTGTACGGTCC---GCCACGCCACCGTCACTTCCAGAGAGAGGAGACAG 822
Qy 893 TGGCGGAGCAGAGAGGAGCGAGCAGCCATGATGTTGGGAATTCGATTTGGCGTGT 952
Db 823 TGGCGGAGCAGAGAGGAGCAGCGGCGCGCTCATGGTGGCATCTCTATTGGTGTGTC 882
Qy 953 GTGCTGTGCTGATCCCTTCTCTGACGGAATCTCATCAGCCCACTCTGTGCTGACG 1012
Db 883 GTGCTGTGCTGATCCCTTCTCTTCTCACCGAGCTCATCAGTCCCTCTGCTCTGAC 942
Qy 1013 CTGCCCCCATCTGGAAGCATATTTCTGTGGCTTGGCTACTCCAAATTTCTTCTCAAC 1072
Db 943 ATCCCCACCATCTGGAAGCATATTTCTGTGGCTTGGCTACTCCAACTCTCTTTAAC 1002
Qy 1073 CCCCTGATTTACAGCTTTTAAACAAGAACTACAACATGCTTCAAGAGCTCTTTTACT 1132
Db 1003 CCCCTGATCTATACGGCTTTCAACAAGAACTACAACAGTGCCTTCAAGAACTATTTCT 1062
Qy 1133 AAGCAGATGA 1144
Db 1063 AGGCAACACTGA 1074

```

```

RESULT 3
LOCUS BZ601879/c 640 bp DNA linear GSS 08-JUN-2003
DEFINITION WHAA55TF Human MCF7 breast cancer cell line library (MCF7_1) Homo
          sapiens genomic clone MCF7_1-1113, genomic survey sequence.
ACCESSION BZ601879
VERSION BZ601879.1 GI:31510341
KEYWORDS GSS.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens

```

```

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 640)
Volik, S., Zhao, S., Chin, K., Brebner, J.H., Herndon, D.R., Tao, Q.,
Kowbel, D., Huang, G., Lapuk, A., Kuo, W.-L., Magrane, G., de Jong, P.,
Gray, J.W. and Collins, C.
End-sequence profiling: Sequence-based analysis of aberrant genomes
Proc. Natl. Acad. Sci. U.S.A. 100 (13), 7696-7701 (2003)
22709111
12788976
Contact: Volik SV
Colin Collins' lab
UCSF Comprehensive Cancer Center
UCSF Box 0809, San Francisco, CA 94143-0808, USA
Tel: 415 502 7066
Fax: 415 502 5665
Email: svolik@cc.ucsf.edu
This clone is available from Amplicon Express
http://www.genomex.com
Class: BAC ends.
Location/Qualifiers
1..640
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/clone="MCF7_1-1113"
/sex="female"
/clone_lib="Human MCF7 breast cancer cell line library
(MCF7_1)"
/note="Vector: pECBAC1; Site 1: HindIII; This library was
constructed from MCF7 breast cancer cell line by Amplicon
Express (http://www.genomex.com) using their standard
procedure."

```

ORIGIN

```

Query Match 41.5%; Score 477.6; DB 28; Length 640;
Best Local Similarity 92.2%; Pred. No. 4.3e-77;
Matches 535; Conservative 0; Mismatches 19; Indels 26; Gaps 2;

Qy 270 GCGATACTTGGTGGCTCGACGGCGCTCGGACGAACTAGTGGCAGCGTGGCGATGC 329
Db 640 GCGATACTTGGTGGCTCGACGGCGCTCGGACGAACTAGTGGCAGCGTGGCGATGC 581
Qy 330 CACGAG--CTTGGCGAGTGAAGTGTGACCGGCGGCACTGCGCTGCGCGGAGCCTG 388
Db 580 CACGAGCCTTGGCGAGTGAAGTGTGACCGGCGGCACTGCGCTGCGCGGAGCCTG 521
Qy 389 TGCCACGTGTGATCTCCTTCGACCGCGGAGCCTGT----- 424
Db 520 TGCCACGTGTGATCTCCTTCGACCGCGGAGCCTGTGCGGAGTGTGATCTCCTTCCAG 461
Qy 425 -CTGTGTGCCCCCGCGGCTCGGGAACGTGGGCGCATCGCCTGTGGCGCGAGCGGCG 483
Db 460 GCTGTGTGCCCCCGCGGCTCGGGAACGTGGGCGCATCGCCTGTGGCGCGAGCGGCG 401
Qy 484 CATCACGCGCACCTGCGAGCACAGCTGCGACCGGAGCGCGCTGTTGCTCATGAT 543
Db 400 CATCACCGCACCTGCGAGCACAGCTGCGACCGGAGCGCGCTGTTGCTCATGAT 341
Qy 544 CGCGCTCGCCCGGGTGGCGTGGCGGCTCATCGCCTGCGCGCTGTCTTTTGGCGGGG 603
Db 340 CGCGCTCACCGGGTGGCGTGGCGGCTCATCGCCTGCGCGCTGTCTTTTGGCGGGG 281
Qy 604 CGAGGTGTGCAAGCTCGCTCGCTCCAGCGCTGCCAGTGAGCGGGAACCTCTTATGCCG 663
Db 280 CGAGGTGTGCAAGCTCGCTCGCTCCAGCGCTGCCAGTGAGCGGGAACCTCTTATGCCG 221
Qy 664 CTTCTCACCGCGCGCTTCCACCTGCGCTTGGCGTGGTGGTGGTGGTGGTGGTGGT 723
Db 220 CTTCTCACCGCGCGCTTCCACCTGCGCTTGGCGTGGTGGTGGTGGTGGTGGTGGT 161
Qy 724 GATCTACGAGCGGCCCAAGTTTCGTTTTCGCGCGCGCGGAGAGCTGTGCTGCGCTTGC 783

```

```
Db 160 GATCTACGAGCGGCGCAAGTTTCGTTTGGCGCGCGGAGAGCTGTGTCGCGTTGCC 101
QY 784 GCCACCATCGAGTGCAGTCCAAAGGTAAGGAAGACACT 823
Db 100 GGCCACCATCGAGTGCAGGTTGGGCTGAGGAACGTTGCT 61

RESULT 4
AY402563 1074 bp DNA linear GSS 15-DEC-2003
LOCUS Mus musculus HTR5A gene, VIRTUAL TRANSCRIPT, partial sequence,
DEFINITION genomic survey sequence.
ACCESSION AY402563
VERSION AY402563.1 GI:39758549
KEYWORDS GSS.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 1074)
Clark, A.G., Gnanowski, S., Nielson, R., Thomas, P., Kejariwal, A.,
Todd, M.A., Tanenbaum, D.M., Civello, D.R., Lu, F., Murphy, B.,
Ferriera, S., Wang, G., Zheng, X.H., White, T.J., Sninsky, J.J.,
Adams, M.D. and Cargill, M.
TITLE Inferring nonneutral evolution from human-chimp-mouse orthologous
gene trios
JOURNAL Science 302 (5652), 1960-1963 (2003)
PUBMED 14671302
REFERENCE 2 (bases 1 to 1074)
Clark, A.G., Gnanowski, S., Nielson, R., Thomas, P., Kejariwal, A.,
Todd, M.A., Tanenbaum, D.M., Civello, D.R., Lu, F., Murphy, B.,
Ferriera, S., Wang, G., Zheng, X.H., White, T.J., Sninsky, J.J.,
Adams, M.D. and Cargill, M.
TITLE Direct Submission
JOURNAL Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive,
Rockville, MD 20850, USA
COMMENT This sequence was made by sequencing genomic exons and ordering
them based on alignment.
FEATURES
source Location/Qualifiers
1..1074
/organism="Mus musculus"
/mol_type="genomic DNA"
/db_xref="taxon:10090"
<1..>1074
/gene="HTR5A"
/locus_tag="HOM1265"
ORIGIN

Query Match 37.0%; Score 426.2; DB 29; Length 1074;
Best Local Similarity 67.4%; Pred. No. 1.2e-67;
Matches 685; Conservative 0; Mismatches 308; Indels 24; Gaps 5;

QY 128 GTCTCGCGGCGGCGGCGGCGGCGGCTTCTCTGTCTTCTACGGTCTCTGTGTGGTGAGCGTGTCTA 187
Db 82 GTCTCGCGGCGGCGGCGGCGGCGGCTTCTCTGTCTTCTACGGTCTCTGTGTGGTGAGCGTGTCTA 141
QY 188 GTCTCGGCTGATCGCTGCACCTTTCCTGTGGAACTGCTGTGCTCGGTCACCATCCCGCGG 247
Db 142 GGCTTTCTAGCTCGGCGCACATTCATCTTGAACTGCTGTGTGGTACCATCTCAAG 201
QY 248 GTCCGCTGCTTCCACCGCGTCCGCATAACTTGGTGGCTTCGACGCGCGTCTCGGACGAA 307
Db 202 GTACGCACCTTCCACCGAGTACCACACAACTGTAGTTCCTATGGCCATCTCGGATGTG 261
QY 308 CTAGTGGCAGCTGGGATGCCACCGAGCTGGCAGTGTGCTGTGCTGACGCGGCGGACGT 367
Db 262 CTAGTGGCTGTGCTGTATTATGCCACTGAGCTGTATCATGAGCTGT---CTGGGCGCGG 318
QY 368 CGCTCTGCTGGGCGGAGCGTGTGCCAGTGTGTGATCTCTTCGACGCGGAGCGTGTGCTG 427
Db 319 TGCAGCTGGGCGGCGGCTGTATGCCAGCTGTGGATCGCATGTGACG-----TGCTC 369
QY 428 TGCTGCCCCCGCGGCTTCGGGAACGTGGCGGCCATGCGCCCTGGGCGCGGAGCGGCGCATC 487
```

```
Db 370 TGCTGTACTGCCAGCATCTGGAATGTGACAGCAATAGCTCTGACCGCTACTGGTCAATC 429
QY 488 ACACGGCAGCTGCAGCAGCAGCTGCGCAGCCGCGCGCGCTGTGCTCATGATCGCG 547
Db 430 ACGCGCCACCTGGAGTACACACTCCGTACCCGCAAGCGTGTCTTCCAATGTATGATCTCTG 489
QY 548 CTCGCCCGGCTGCGGCGGCTCATCGCCCTCGCGCGCTGCTCTTTGGCGGCGGCGAG 607
Db 490 CTCACCTGGGCACTCTCCACTGTCTCTCTGGCTCCACTGCTATTTGGTGGGAGAG 549
QY 608 GTGTGGAGCCTCGGCTCCAGCGCTGCGAGGTGAGCCGGGAACCTCTCTATGCCGCTTC 667
Db 550 ACTTATTCTAGCCCCAGTGCAGGAATGCCAAGTCAAGTCAGTCGCGAGCTTCTTACACCGTGTTC 609
QY 668 TCCACCCGCGGCGGCTTCCACCTGCGCTTGGCGTGTGCGCTGCTCTTCTACCGGAAGATC 727
Db 610 TCCACCGTGGGCTTCTTACCTGCGCTGTGCGTGTGCTCTTGTGTACTGGAATTT 669
QY 728 TAGAGGCGGCGCAAGTTTCGTTTGGCCCGCGCGAGAGCTGTGCTGCGGTTGCCGCG 787
Db 670 TACAGGCGCGCAATTTCCGATGGGCTCCAG--GAAGACCAACACAGCGTCTCCCGCTAC 727
QY 788 ACATGCGAGTGAAGTCCAAAGTAAGGAAGCACCTGATGAGGCTGAAGTGTGTTCACG 847
Db 728 CCGAAGCTGTGG-----AGGTGAAGAATGCTTACACAACATCCCCAGATGGTGTTCACG 780
QY 848 GCACATTGCAAGCAACGAGTGTCTTCCAGGTGAGCGGGGACTCTCTGGCGGAGCAGAAG 907
Db 781 GTCC---GCCATGCCACCGCTCACCTTCCAGAGAAGGGATACGTGGAGGAGCAGAAG 837
QY 908 GAGAGCGAGCAGCAGCATGATGTGGAAATCTGATTTGGCGTGTGTGTGCTGTGCTGCATC 967
Db 838 GAGCAAGGCGCAGCCCTCATGTGGGCATCCTCATCGAGTGTGTGTGCTGTCTGCTGCTC 897
QY 968 CCCTTCTTCTGACGGAATCATCAGCCACTCTGTGCTGAGCTGAGCTGCCCCCATCTGG 1027
Db 898 CTTTCTTCTGTCACAGAGCTCATGTCCTGCTGTGTTCTCTGGAGCTGCTGCTGCTGCTG 957
QY 1028 AAGAAGCATATTTCTGTGGCTTGCTACTCCTCAATCTTCTTCAACCCCTGATTACACA 1087
Db 958 AAGAGCATCTTCTGTGGTGGGCTATTCTTAATCTTCTTCAACCCATCATCTACACA 1017
QY 1088 GCTTTTAAACAAGAACTACAAATGCTTCAAGAGCTCTTTACTAAGCAGAGATGA 1144
Db 1018 GCATTCAACAGGAGCTACAGAGCTGTCTTCAAGGCTCTTCTTCTCCAAGCAACAATGA 1074

RESULT 5
BM805859 1099 bp mRNA linear EST 05-MAR-2002
LOCUS AGENCOURT_6513650 NIH_MGC_124 Homo sapiens cDNA clone IMAGE:5731557
DEFINITION S', mRNA sequence.
ACCESSION BM805859
VERSION BM805859.1 GI:19122682
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 1099)
NIH-MGC http://mgi.nci.nih.gov/.
AUTHORS National Institutes of Health, Mammalian Gene Collection (MGC)
TITLE Unpublished (1999)
JOURNAL Contact: Robert Strausberg, Ph.D.
COMMENT Email: cgabbs-r@mail.nih.gov
Tissue Procurement: Invitrogen
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LNL at:
http://image.lnl.gov
```


Plate: LLAM12732 row: d column: 22
High quality sequence start: 7
High quality sequence stop: 616.
Location/Qualifiers
1. 1099
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:5731557"
/tissue_type="hippocampus"
/lab_host="DH10B"
/clone_lib="NIH_MGC 124"
/note="Organ: brain; Vector: pCMV-SPORT6; Site 1: EcorV (destroyed); Site 2: NotI; RNA source male hippocampus, age 27. Library is oligo-dr primed and directionally cloned (EcorV site is destroyed upon cloning). Average insert size 1.4 kb, insert size range 0.9-4 kb. Library is normalized and enriched for full-length clones and was constructed by C. Gruber (Invitrogen). Research Genetics tracking code 012."

ORIGIN

Query Match 24.7%; Score 284.2; DB 12; Length 1099;
Best Local Similarity 71.2%; Pred. No. 1.1e-41;
Matches 408; Conservative 0; Mismatches 153; Indels 12; Gaps 2;
Qy 572 ATCCCTCGGCGCGCTCTTTGGCGGGGAGGTGTCGAGCTCGGCTCCAGGC 631
Db 18 ATCTCTGCGCGCGCTCTTTGGCTGGGAGAGACGTACTCTGAGGCGAGCGAG 77
Qy 632 TGCCAGGTGAGCGGGACCTCTATGCCGCTTCTCCACCGGGCGCTTCCACCTG 691
Db 78 TGCCAGGTGAGCGGGACCTCTTACGCCGTGTTCTCCACGTTAGGCGCTTCTACCTG 137
Qy 692 CCGCTTGCGCGTGGTCTTCTACCGAAGATCTACGAGGCGGCGCAAGTTGTTTC 751
Db 138 CCGCTTGTTGGTCTCTTCTGTTGTTGGAAGATCTCAAGGTGCGCAAGTTCCGCTG 197
Qy 752 GCGCGCGCGGAGAGCTGTCTGCTGCGGTCACCATGTCAGGTGAGGTCCAGGTA 811
Db 198 GGTCTCCAGGAAGACCAATAGCGT-----CTCACCCATATCCGAAGCTGTGGAGGTG 248
Qy 812 AAGGAAGCACTGATGAGGTGTAAGTGTGTTTACGGGCACATTCGAAAGCAAGCTGTC 871
Db 249 AAGGACTCTGCAACAGACCCAGTGTGTTTACGGTCC---GCCAGCCACCGTCAAC 305
Qy 872 TTCCAGGTGAGCGGGACTCTTGGCGGAGCAGAAAGAGAGCGAGCGCATGATGTTG 931
Db 306 TTCCAGCCAGAAAGGACACGTGGCGGAGCAGAAAGAGAGCAGCGCGCGCTCATGTTG 365
Qy 932 GGAATTCGATTGGGTGTTGTGTGCTGTGGATCCCTTCTTCTGACGGAACCTCATC 991
Db 366 GGCATCTCATGGCGTGTGTGTCTGTGATCCCTTCTTCTCACCAGTCTCATC 425
Qy 992 AGCCCACTGTGCTGCGAGCTGCCCCCACTCGAAAGATATTTCTGTGGCTTGGC 1051
Db 426 AGTCCCTCTGCTCTGTGACATCCCGGCATCTGGAAGATCTTCTGTGGCTTGGC 485
Qy 1052 TACTCCAATCTTTCTTCAACCCCTGATTTTACAGCTTTTAAAGAACTTAAACAAAT 1111
Db 486 TACTCCAATCTTCTTTAAACCCCTGATCTATACGGCTTTCAACAAGAACTTAAACAGC 545
Qy 1112 GCTTCAAGAGCTCTTTACTAAGCAGATCA 1144
Db 546 GCTTCAAGAACTTCTTTTCTAGGCAACACTGA 578

RESULT 6

BI489745
LOCUS 906 bp mRNA linear EST 28-AUG-2001
DEFINITION 603032245F1 NIH_MGC_115 Homo sapiens cDNA clone IMAGE:5173354 5',
mRNA sequence.
ACCESSION BI489745

VERSION
KEYWORDS
SOURCE
ORGANISM

BI489745.1 GI:15328973
EST.
Homo sapiens (human)
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
NIH-MGC http://mgs.nci.nih.gov/.
1 (bases 1 to 906)
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgaps-remail.nih.gov
Tissue Procurement: Life Technologies, Inc.
cDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM11431 row: j column: 11
High quality sequence stop: 826.
Location/Qualifiers

FEATURES

source

1. 906
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:5173354"
/lab_host="DH10B"
/clone_lib="NIH_MGC_115"
/note="Organ: pooled brain, lung, testis; Vector: pCMV-SPORT6; Site 1: NotI; Site 2: EcorV (destroyed); RNA source anonymous pool of 6 male brains, age range 23-27; 1 male lung, age 27; and 1 male testis, age 69. Library is oligo-dr primed and directionally cloned (EcorV site is destroyed upon cloning). Average insert size 1.8 kb, insert size range 1-3 kb. Library is normalized and enriched for full-length clones and was constructed by C. Gruber (Invitrogen). Research Genetics tracking code 021. Note: this is a NIH_MGC Library."

ORIGIN

Query Match 24.3%; Score 280.4; DB 12; Length 906;
Best Local Similarity 71.1%; Pred. No. 5e-41;
Matches 431; Conservative 0; Mismatches 161; Indels 14; Gaps 4;
Qy 540 TGATCGCGCTCGCGGTCGCGGCTCATCGCCTCGCGCGCTGCTCTTTGGCC 599
Db 1 TGATCGCGCTCACCTGGGCACTCTCCGCTGTCTCTGCGCCCGCTGCT- 59
Qy 600 GGGGCGAGGTGTGGACGCTCGGCTTCCAGCGCTGCCAGGTGAGCGGGAACCTCTATG 659
Db 60 GGGGAGAGACGTACTCTGAGGCGAGGAGTGCCAGGTAAGCGCGAGCCTTCTACG 119
Qy 660 CCGCTTCTCCACCGCGCGCTTCCACCTGCGCTTGGCGTGTGCGGTTGTCTACC 719
Db 120 CCGTGTCTCCACCGTAGCGCTTCTACCTGCGCTCTGTGTGTCTCTCGGTACT 179
Qy 720 GGAATCTACGAGCGCGCAAGTTTCGTTTCGCGCGCGCGAGAGCTGTGTGCGCTG 779
Db 180 GGAAGATCTCAAGGCTCCAGGTTCCGCTGGGTCCAGGAAGACCAATACGCT- 234
Qy 780 TGCCGCGCCACCATTCAGGTGAGGTCCAAAGTAAAGAAAGACCTTGATGAGGTGAAGTGG 839
Db 235 -----CTCACCATATCCGAAGCTGTGGAGGTCAAGACTCTGCCAAACAGCCAGATGG 290
Qy 840 TGTTCAAGGCAATTCGAAAGCAAGGTGCTCTCCAGGTGAGCGGGGACTCTCGCGGG 899
Db 291 TGTTCAAGGTCC---GCCACGCCACCGTCACCTTCCAGCCAGAGGGGACACGTGCGGG 347
Qy 900 AGCAGAGGAGCGGAGCGGACCATGATGTTGGGAACTTCTGATTCGCGTGTGTTGCTGT 959
Db 348 AGCAGAGGAGCGGCGCGCCCTCATGGTGGGCACTCTCATTTGGGCTGTTTCGCTCT 407

Qy	960	GCTGGATCCCTTCTTCTCTGACGGAACTCATCAGCCCACTCTGTCGCTGAGGCTGCCCC	1019
Db	408	GCTGGATCCCTTCTTCTTCTCACCAGCTCATCAGTCCCTCTGCTCTGTGACATCCCG	467
Qy	1020	CCATCTGGAAAAGCATATTTCTGGGCTTGGCTACTCCAATCTTTCTTCAAACCCCTGA	1079
Db	468	CCATCTGGAAAAGCATCTTCTGTGGCTTGGCTACTCCAATCTTCTTTAAACCCCTGA	527
Qy	1080	TTTACACAGCTTTTAAC-AAGAACTCAACAAATGCCTTCAAGGCTCTTTTACTAAGCAG	1138
Db	528	TCTATACGGCTTTCAACGAAGAACTACACAGGCCCTTCAAGAACTTCTTTCTTAGGAA	587
Qy	1139	AGATGA	1144
Db	588	CACTGA	593

RESULT 7					
EX280849					
LOCUS	EX280849	539 bp	mRNA	linear	EST 04-MAR-2003
DEFINITION	EX280849	NIH MGC 115	Homo sapiens	cDNA clone IMAGE:5173354	IMAG:5173354, mRNA sequence.

ORIGIN

Query Match 23.0%; Score 265.4; DB 13; Length 539;
Best Local Similarity 70.4%; Pred. NO. 2.4e-38;
Matches 388; Conservative 0; Mismatches 151; Indels 12; Gaps 2;

QY	540	TGATCGCGCTGCGCCGGTGCCTGCGCGCTCATGCGCCCTGCGCGCGTCTCTTTGGCC	599
Db	1	TGATCGCGCTCACTGGGCACTCTCGCTGTCACTCTCTGGCCCCGCTGCTTTTGGCT	60
QY	600	GGGCGAGGTGTGGAGCTCGGCTCCAGCGCTCCAGGTGACGCGGGAACCTCTATG	659
Db	61	GGGAGAGAGCTACTCTGAGGGAGGAGGTGCCAGGTAAGCCCGAGCTTCCTACG	120
QY	660	CCGCGTTCTCCACCCCGCGCGCTTCACCTGCGGCTTGCGGTGTCGCGTTTGTCTACC	719
Db	121	CCGTGTTCTCCACCGTAGGCGCTTCTACTCTGCGCTCTGTGTGTGCTCTTCGTGTACT	180
QY	720	GGAAGATCTACGAGGCGGCGAAGTTTCGTTTCGCGCGCGCGAGAGCTGTGCTGCCGT	779
Db	181	GGAAGATCTACAAGGCTGCCAAGTTCGCGTGGGCTCCAGGAAGACCAATAGCGT	235
QY	780	TGCGGCGCACATGCAGGTGAGTCCAGGTAAGGAAGCACTGTATGAGGTGAAGTGG	839
Db	236	-----CTCACCCTATCTCGAAGCTGTGGAGGTGAAGACTCTGCAACAGCCCGCAGTGG	291
QY	840	TGTTACGGCACATTGCAAGCAACGGTGTCTTCAGGTGAGCGGGGACTCTCTGGCGGG	899
Db	292	TGTTACGGTCC--GCCAGCCACCGTCACCTTCAGCCAGNAGGGGACAGTGGCGGG	348
QY	900	AGCAGAGGAGAGCGGAGAGCCATGATGTGGAAATCTGATGTGCGTGTGTGTCTGT	959
Db	349	AGCAGAGGAGAGCGGGCGCGCTCATGTGGGCATCCTCATTTGGCGTGTTCGTGCTCT	408
QY	960	GCTGGATCCCGTCTTCTTCAGGAACTCATAGCCCACTCTGTGCTTCAGCGCTGCC	1019
Db	409	GCTGGATCCCGTCTTCTTCACCGAGCTCATCAGTCCCTCTGCTCTGTGACATCCCCG	468
QY	1020	CCATCTGGAAAACATATTCTGTGGCTTGGGTACTTCCAAATCTTCTTCAACCCCGCTGA	1079
Db	469	CCATCTGGAAAACATCTTCTGTGGCTTGGGTACTTCCAACTCTTCTTAAACCCCTGA	528
QY	1080	TTTACAGCT	1090
Db	529	TCATACGGCT	539

RESULT 8	
CNS05MTV	
LOCUS	CNS05MTV 1027 bp DNA linear GSS 01-SEP-2000
DEFINITION	Tetraodon nigroviridis genome survey sequence T3 end of clone 038018 of library A from Tetraodon nigroviridis, genomic survey sequence.

[illegible]

RESULT 9
BI603546/c


```

source
1. .583
/organism="Bos taurus"
/mol_type="mRNA"
/db_xref="taxon:9913"
/tissue_type="pooled"
/lab_host="DH10B"
/clone_lib="MARC 480V"
/notes="vector: pCMV SPORT6; Site 1: NotI; Site 2: SalI;
Library made from pooled tissue from day 20 and day 40
embryos."

ORIGIN

Query Match      21.5%; Score 247.6; DB 12; Length 583;
Best Local Similarity 68.0%; Pred. No. 4.3e-35;
Matches 413; Conservative 0; Mismatches 170; Indels 24; Gaps 4;

QY 322 GCGGATGCCACCGAGCCTGCGAGTGTGATCTCTCTTCAGCCGCGAGCTGTGTGTCTGCTCCCGCCGCG 381
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1 GGTGATGCCCTGAGCCTGTGTGACGAGCTGT---CGGGCGCGCTGGCAGCTGGGCGG 57
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 382 GAGCCTGTGCCACAGTGTGGATCTCTCTTCAGCCGCGAGCTGTGTGTCTGCTCCCGCCGCG 441
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 58 GCGGCTGTGCCAGTGTGGATCGGTGCGATGTG-----CTCTGCTCAGCGCCAG 108
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 442 CTTGCGGAAAGTGTGGCGGCATTCGCCCTGGGGCGGAGCGGGGCCATCAACGCGACCTGCA 501
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 109 CATCTGAAGCTGACGGCCATTCGGCTGGACCGCTACTGTGTCCATCACCGCTCAGCTGGA 168
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 502 GCACAGCTCGGCACCGCGAGCGGCGCTGTTGCTCATGATCGGCTCGCCCGCGGTGCC 561
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 169 ATACACACTCGGCGCGCGAGCGGCTCTCCAAAGTCATGATCGGCTCACCTGGGCGCT 228
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 562 GTCCGCGCTCATTCGCCCTCGCGCGCTGCTCTTTGGCGGGCGAGGTGCGACGCTCG 621
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 229 CTCGCCGTTCATCTCTTTGCCCCACTGCTCTTCGGCTGGGGGAGACCTACTCAGAGG 288
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 622 GCTTCAGCGTTCGAGTGAGCGGGAAACCTCTCTATGCCGCTTCTCCACCGCGGCGC 681
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 289 CAGCGAGGAGTGCCAGGTGAGCGCGAGCCCTCTACACCGTGTCTCACCGTGGGCGC 348
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 682 CTTCCACTGCGCTTGGCGTGTGCGGTTGTCTACCGGAAGATCTACGAGAGCGGCCAA 741
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 349 CTTCTACTCGCGTGGCGGTGTGCTTTGTGTACTGGAAGATCTACAGGCGCGCAA 408
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 742 GTTTCGTTTCGGCGCGCGCGAGAGCTGTCTGCGGTGTCGGCCACCATTCAGGTGAG 801
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 409 GTCGCGGTGGGCCCCCGGAAGACCAACAGCGT-----CTCACCATATCCGAAC 459
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 802 GTCCAAAGTAAAGGAAGACCTGATAGGCTGAAGTGTGTTCACGGCACATTCGAAAGC 861
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 460 CGTGACGGGAAGGCGCGGCTCAGCGGCTCAGATGGTCTCGCGGCGC---GCCGCGC 516
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 862 AACGGTCTCTTCAGGTGACGGGACTCTCTGCGGAGCAGAGGAGCGGAGCGAGCAGC 921
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 517 CAAGGTCACTCTTCAGGCGGAGCGGTGCGGAGCAGAGGAGCGGCGCGC 576
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 922 CATGATG 928
    ||| ||| |||
Db 577 CTTATG 583
    ||| ||| |||

RESULT 12
BG715343/c
LOCUS BG715343
DEFINITION 602677363F1 NIH_MGC_96 Homo sapiens cdna clone IMAGE:4799726 5',
mRNA sequence.
ACCESSION BG715343
VERSION BG715343.1
KEYWORDS GI:13994526
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

```

```

RESULT 13
BI411238
LOCUS      929 bp      mRNA      linear      EST 14-AUG-2001
DEFINITION 602964578F1 NCI_CGAP_Lu33 Mus musculus cDNA clone IMAGE:5120035 5',
mRNA sequence.
ACCESSION  BI411238
VERSION     BI411238
KEYWORDS    EST.
SOURCE      Mus musculus (house mouse)
ORGANISM    Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 929)
NIH-MGC http://mgc.nci.nih.gov/.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: Gilbert Smith, Ph.D.
cDNA Library Preparation: M. Bento Soares, Ph.D., M. Fatima
Bonaldo, Ph.D.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: NCI-CGAP clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLNL1292 row: 1 column: 20
High quality sequence start: 177
High quality sequence stop: 928.

FEATURES             source
1..929
    Location/Qualifiers
        /organism="Mus musculus"
        /mol_type="mRNA"
        /strain="Czech II"
        /db_xref="taxon:10090"
        /clone="IMAGE:5120035"
        /tissue_type="pooled lung tumors"
        /lab_host="DH10B (phage-resistant)"
        /clone_lib="NCI CGAP_Lu33"
        /note="Organ: lung; Vector: pTTT3D-Pac (Pharmacia) with a
        modified polylinker; Site 1: NotI; Site 2: EcoRI; 1st
        strand cDNA was prepared from mRNA obtained from pooled
        lung tumors with a Not I - oligo(dT) primer [5'
        TGTTCACCAATCTGAAGTGGAGCGCGCTCTGTCTTTTCTTTTCTTTT 3'].
        Double-stranded cDNA was ligated to Eco RI adaptors
        (Pharmacia), digested with Not I and cloned into the Not
        I and Eco RI sites of the modified pTT3 vector. Library
        went through one round of normalization, and was
        constructed by Bento Soares and M. Fatima Bonaldo. "

ORIGIN
Query Match      17.4%; Score 200.4; DB 12; Length 929;
Best Local Similarity 69.5%; Pred. No. 2.1e-26;
Matches 421; Conservative 0; Mismatches 161; Indels 24; Gaps 10;

QY      5  ATGGAGGCGCTAGCCTTTTCAGTGGCCACCGCGCGGTTCGCCCTTGGCCCTGGGACCCGAG 64
      |||||
Db      338  ATGGAAGTTTCTAACTCTCAGCGGCCACTCCCGGACTTGCCTTCTCTCCGGACCC--- 394
      |||||

QY      65  ACCACAGCGGACCGGGGACCCACGCGGAGGAGTACTCGTTTCGACCCCGAGCGGC 124
      |||||
Db      395  -TGAGAGTTCGAGTGACAGCCCAAGTTCCGTGCAGGAGCATGGGATCCACCCCGAGTGGG 453
      |||||

QY      125  GCGGTCCTGCGCGGGCGG--AGGGCGCCCTTCTGTCTTCA-CGTCCTGTGTGTGAGC 181
      |||||
Db      454  CTCATCTTGCCTGTGTCGGAGCGCTGAATCTTCTGCTTTCACCTGTGCTGTGTGTGACT 513
      |||||

QY      182  CTGCTAGTATGC--TGCTGATCG-CTGCCACTTTCCTGTGGAACTGCTGTTCCCGTCCACC 238
      |||||
Db      514  CTACTTGGTGTGTTGCTGATGCTTTGGCACTTCTTATGGAATCTGCTAGTCTGTGACT 573
      |||||

QY      239  ATCCCGCGGGTCGGTGCCTTCACCGCGTCCGCACTACTTGGTGGCCTCGAGCGCGCTC 298
      |||||

```

```

Db      574  ATCTTGGGTGTCGGGCGCTTCCACCGGTGCCACATAACTTGGTGGCTCGACAGCGCTC 633
      |||||
QY      299  TCGGACGAACCTAGTGGCAGCGCTGGGATGCCACCGAGCTGGGAGTGGCTGTGCAGC 358
      |||||
Db      634  TCGGATGTCTCTGGTGGCGCTCTGGTGATGCGCTCTGAGCTGGTGGAGAGTGTCCGT- 692
      |||||
QY      359  GGGCGACGTCGGCTGCTGGGCGGAGCCTGTGCCAGTGTGGATCTCTTCCTCCAGCGCGGA 418
      |||||
Db      693  GGGCGACGTTGGCAGCTAGGAGGAGTCTGTGCCAGTGTGGATCTCTCTTCGACGTGTAG 752
      |||||
QY      419  GGCTGTGTGTGTGCCCCCGCGCTCGGGAACGTGGCGGCATCGCTTGGCCCGCGAC 478
      |||||
Db      753  TGTCTG-----CACCGCAGCATTTGGAACGTGGCTGGCATCGCTTGGATCGCTAC 802
      |||||
QY      479  GGGGGCATCACAGCGCACCTGGCAGCACAGCTGC-GCACCCGACGCGCGCTT-CGTTGC 536
      |||||
Db      803  TGGACTATCACGGGCA-CTGCAGTACAGCGTGTGCACCCGGAGCGGTCTACTGCGC 861
      |||||
QY      537  TCATGATCGCGCTCGCCCGGCTGCGGTGCGCGCTCATCGCCCTCGCGCGCGCTCTTTG 596
      |||||
Db      862  TCATGATCGGATCACCTGGGCACTGTCCGGCGCTCATTTGCTTCGCGCGCTGCTTAGS 921
      |||||
QY      597  GCCGGG 602
      |||||
Db      922  CTGGGG 927

RESULT 14
BB628489
LOCUS      660 bp      mRNA      linear      EST 26-OCT-2001
DEFINITION BB628489 RIKEN full-length enriched, 16 days neonate cerebellum Mus
musculus cDNA clone 9630009M09 5', mRNA sequence.
ACCESSION  BB628489
VERSION     BB628489.1 GI:16465907
KEYWORDS    EST.
SOURCE      Mus musculus (house mouse)
ORGANISM    Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 660)
Arakawa,T., Carninci,P., Fukuda,S., Furuno,M., Hanagaki,T.,
Hara,A., Hiramoto,K., Hori,F., Ishii,Y., Ito,M., Kawai,J.,
Konno,H., Kouda,M., Koya,S., Matsuyama,T., Miyazaki,A., Nomura,K.,
Ohno,M., Okazaki,Y., Okido,T., Saito,R., Sakai,C., Sakai,K.,
Sano,H., Sasaki,D., Shibata,K., Shinagawa,A., Shiraki,T.,
Sogabe,Y., Suzuki,H., Tagami,M., Tagawa,A., Takahashi,F.,
Takeda,Y., Tanaka,T., Toya,T., Muramatsu,M. and Hayashizaki,Y.
RIKEN Mouse ESTs (Arakawa,T., et al. 2001)
Unpublished (2001)
Contact: Yoshihide Hayashizaki
Laboratory for Genome Exploration Research Group, RIKEN Genomic
Sciences Center(GSC), Yokohama Institute
The Institute of Physical and Chemical Research (RIKEN)
1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
Tel: 81-45-503-9222
Fax: 81-45-503-9216
Email: genome-res@gsc.riken.go.jp
URL:http://genome.gsc.riken.go.jp/
Carninci,P., Shibata,Y., Hayatsu,N., Sugahara,Y., Shibata,K.,
Itoh,M., Konno,H., Okazaki,Y., Muramatsu,M. and Hayashizaki,Y.
Normalization and subtraction of cap-trapper-selected cDNAs to
prepare full-length cDNA libraries for rapid discovery of new
genes. Genome Res. 10 (10), 1617-1630 (2000)
wagi,K., Fujiwaka,S., Inoue,K., Togawa,Y., Izawa,M., Ohara,E.,
Watanabe,M., Yoneda,Y., Ishikawa,T., Ozawa,T., Tanaka,T.,
Matsura,S., Kawai,J., Okazaki,Y., Muramatsu,M., Inoue,Y., Kira,A.
and Hayashizaki,Y.
RIKEN integrated sequence analysis (RISA) system--384-format
sequencing pipeline with 384 multicapillary sequencer. Genome Res.
10 (11), 1757-1771 (2000)
Konno,H., Fukunishi,Y., Shibata,K., Itoh,M., Carninci,P.,
Sugahara,Y. and Hayashizaki,Y.

```


Db 655 CCCCCAACACTTGTATGACCTCCATGGCCATATCTGACGTGATGCTGAGCTTTGGTGATG 596
QY 329 CCACCGAGCTGCGGAGTGAGCTGTGACCGGGCGACGTGCGCTGCTGGCGCGAGCCTG 388
Db 595 CCTCTCAGTCTGTGTCACAGCT---GAACGGCCGTGCGTGAAGCTGGTCCGTTCTC 539
QY 389 TGCACGCTGTGGATCTCTTCGACGCGGAGCCTGTCTGTGCTGCCCGCGCCCTCGGG 448
Db 538 TGTCAAGTCTGGATATCGTTGATGTC-----CTGTGCTGTAGGCTAGCAATTGG 488
QY 449 AACGTGGCGCCATCGCCCTGGSCCGGACGGGGCCATCACAGGCACCTGCGAGCACG 508
Db 487 AACGTGACCGCAATAGCCCTCGACCGTTACTGTGTCATTAACCTGACACCTGGAGTACAC 428
QY 509 CTGCGCACCGCGAGCGCGCTCTGTTGCTCATGATCGCGCTGCGCGGGTGCGGTCGGG 568
Db 427 CTAAGACCGGAGAGATCTCCAATGTGATGATCGGGTTAACCTGGCTGCTTTCATCT 368
QY 569 CTCATCGCCCTCGCGCGCTGCTCTTTGGCGGGGCGAGGTGTGCGACGCTCGGCTCCAG 628
Db 367 GTTATTTGCTCTCGCCCTCTG---TTTGGCTGGGTGAGACGTATTTCGGAGGAGAACATG 311
QY 629 CGCTGCCAGTGAGCGGGAACCTCTCTATGCGCCCTTCTCCACCGCGCGCCCTTCCAC 688
Db 310 GAGTGCCAGGTGAGCCAGAGCGCTCTTACACCATCTTTTCCACATTCGGGGCCCTTCTAT 251
QY 689 CTGCGCCTTGGCGTGTGCGCTTGTCTACCGGAAGATCTACGAGCGGCGCAAGTTTCGT 748
Db 250 CTGCGCTCTGTGTGTGTGTGCTTTGTCTACTGGAAGATTACAAAGCTGCTAAGTTCCGT 191
QY 749 TTCGG 753
Db 190 ATTGG 186

Search completed: September 7, 2004, 17:13:12
Job time : 2870 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 7, 2004, 16:25:27 ; Search time 586 Seconds
(without alignments)
9771.633 Million cell updates/sec

Title: US-09-976-782-15

Perfect score: 1152

Sequence:

1 cgccatggaggccgtagcc.....aagcagagatgaacacagg 1152

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 3267054 seqs, 2485319735 residues

Total number of hits satisfying chosen parameters: 6534108

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database :

Published Applications NA:*

- 1: /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq:*
- 2: /cgn2_6/ptodata/2/pubpna/PCT_NEW_PUB.seq:*
- 3: /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq:*
- 4: /cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq:*
- 5: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq:*
- 6: /cgn2_6/ptodata/2/pubpna/PCTUS_PUBCOMB.seq:*
- 7: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq:*
- 8: /cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq:*
- 9: /cgn2_6/ptodata/2/pubpna/US09A_PUBCOMB.seq:*
- 10: /cgn2_6/ptodata/2/pubpna/US09B_PUBCOMB.seq:*
- 11: /cgn2_6/ptodata/2/pubpna/US09C_PUBCOMB.seq:*
- 12: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq:*
- 13: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq:*
- 14: /cgn2_6/ptodata/2/pubpna/US10A_PUBCOMB.seq:*
- 15: /cgn2_6/ptodata/2/pubpna/US10B_PUBCOMB.seq:*
- 16: /cgn2_6/ptodata/2/pubpna/US10C_PUBCOMB.seq:*
- 17: /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq:*
- 18: /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq:*
- 19: /cgn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1152	100.0	1152	10	US-09-976-782-15
2	1068	92.7	1150	10	US-09-823-187-9
3	1064.8	92.4	1150	10	US-09-823-187-11
4	1064	92.4	1458	17	US-10-333-946-34
5	1060	92.0	1119	12	US-10-311-671-25
6	1016	88.2	1155	10	US-09-954-342-29
7	756.2	65.6	3086	15	US-10-017-161-697
8	756.2	65.6	3086	16	US-10-292-798-609
9	732.8	63.6	2061	15	US-10-109-532A-1
10	497.6	43.2	1074	15	US-10-225-567A-445
11	497.6	43.2	1074	15	US-10-345-680-39
12	497.6	43.2	1159	15	US-10-345-680-37
13	497.6	43.2	1159	16	US-10-352-684A-15
14	496	43.1	2700	13	US-09-823-245A-617

15 492.8 42.8 1074 11 US-09-826-509-446
 16 378.2 32.8 636 9 US-09-750-373-5
 17 350.4 30.4 568 15 US-10-029-386-9467
 18 297 25.8 297 9 US-09-750-373-4
 19 275.2 23.9 534 15 US-10-029-386-2940
 20 224 19.4 224 15 US-10-029-386-23167
 21 160.2 13.9 293 15 US-10-029-386-16640
 22 157.6 13.7 1210 13 US-10-027-632-213298
 23 157.6 13.7 1210 16 US-10-027-632-213298
 24 157.2 13.6 585 13 US-10-027-632-213299
 25 157.2 13.6 585 16 US-10-027-632-213299
 26 142.4 12.4 200 15 US-10-109-532A-4
 27 106 9.2 2826 15 US-10-225-567A-43
 28 104.4 9.1 1382 15 US-10-101-510-754
 29 104.4 9.1 1382 16 US-10-305-720-1256
 30 102.8 8.9 1374 15 US-10-001-073-42
 31 102.8 8.9 1386 15 US-10-001-073-40
 32 95.2 8.3 1323 15 US-10-345-680-42
 33 95.2 8.3 1984 15 US-10-225-567A-19
 34 95.2 8.3 1984 15 US-10-101-510-52
 35 95.2 8.3 1984 15 US-10-345-680-40
 36 95.2 8.3 1984 16 US-10-305-720-1077
 37 94 8.2 1332 16 US-10-352-684A-1
 38 94 8.2 1756 16 US-10-305-720-1197
 39 94 8.2 2625 9 US-09-954-531-995
 40 94 8.2 2625 17 US-10-717-597-32
 41 93.6 8.1 1323 11 US-09-826-509-448
 42 93.6 8.1 1356 9 US-09-993-844-12
 43 93 8.1 2428 15 US-10-241-313-15
 44 93 8.1 2455 15 US-10-060-795B-1
 45 93 8.1 2482 15 US-10-060-795B-10

ALIGNMENTS

RESULT 1

US-09-976-782-15
 ; Sequence 15, Application US/09976782
 ; Publication No. US20030190715A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Grosse et al
 ; TITLE OF INVENTION: No. US20030190715A1el Proteins and Nucleic Acids Encoding Same
 ; FILE REFERENCE: 21402-157
 ; CURRENT APPLICATION NUMBER: US/09/976,782
 ; CURRENT FILING DATE: 2001-10-12
 ; PRIOR APPLICATION NUMBER: 60/240,113
 ; PRIOR FILING DATE: 2000-10-12
 ; PRIOR APPLICATION NUMBER: 60/240,662
 ; PRIOR FILING DATE: 2000-10-16
 ; PRIOR APPLICATION NUMBER: 60/240,732
 ; PRIOR FILING DATE: 2000-10-16
 ; PRIOR APPLICATION NUMBER: 60/240,625
 ; PRIOR FILING DATE: 2000-10-16
 ; PRIOR APPLICATION NUMBER: 60/240,703
 ; PRIOR FILING DATE: 2000-10-16
 ; PRIOR APPLICATION NUMBER: 60/241,190
 ; PRIOR FILING DATE: 2000-10-16
 ; PRIOR APPLICATION NUMBER: 60/240,637
 ; PRIOR FILING DATE: 2000-10-16
 ; PRIOR APPLICATION NUMBER: 60/240,669
 ; PRIOR FILING DATE: 2000-10-16
 ; PRIOR APPLICATION NUMBER: 60/262,455
 ; PRIOR FILING DATE: 2001-01-18
 ; PRIOR APPLICATION NUMBER: 60/240,648
 ; PRIOR FILING DATE: 2000-10-16
 ; NUMBER OF SEQ ID NOS: 127
 ; SOFTWARE: PatentIn Ver. 2.1
 ; SEQ ID NO 15
 ; LENGTH: 1152
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 US-09-976-782-15

Sequence 446, App
 Sequence 5, Appli
 Sequence 9467, Ap
 Sequence 4, Appli
 Sequence 23167, A
 Sequence 2940, Ap
 Sequence 16640, A
 Sequence 213298,
 Sequence 213299,
 Sequence 213299,
 Sequence 4, Appli
 Sequence 43, Appli
 Sequence 754, App
 Sequence 1256, Ap
 Sequence 42, Appl
 Sequence 40, Appl
 Sequence 42, Appl
 Sequence 19, Appl
 Sequence 52, Appl
 Sequence 40, Appl
 Sequence 1077, Ap
 Sequence 1, Appli
 Sequence 1197, Ap
 Sequence 995, App
 Sequence 32, Appl
 Sequence 448, App
 Sequence 15, Appl
 Sequence 1, Appli
 Sequence 10, Appl

Db 254 CCGCGGCTCGTGCCTTCCACCGCTGCCGATACCTTGTGGCTCGACGGCCGTCTC 313
QY 301 GGACGAATAGTGGCAGCGCTGGCGATGCGACCGAGCTTGGCGAGTGTGATCGACCGG 360
Db 314 GGACGAATAGTGGCAGCGCTGGCGATGCGACCGAGCTTGGCGAGTGTGATCGACCGG 373
QY 361 GCGAGCTGGCTGGCGCGAGCGCTTGGCGAGCTTGGCGAGCTTGGCGAGCTTGGCGAGC 420
Db 374 GCGAGCTGGCTGGCGCGAGCGCTTGGCGAGCTTGGCGAGCTTGGCGAGCTTGGCGAGC 428
QY 421 CTGTCTGTGCTCCCGCGCGCTTGGCGAGCTTGGCGAGCTTGGCGAGCTTGGCGAGC 480
Db 429 ---CTGTGCTCCCGCGCGCTTGGCGAGCTTGGCGAGCTTGGCGAGCTTGGCGAGC 484
QY 481 GCGCATACACGGGACTTGGCGAGCTTGGCGAGCTTGGCGAGCTTGGCGAGCTTGGCGAGC 540
Db 485 GCGCATACACGGGACTTGGCGAGCTTGGCGAGCTTGGCGAGCTTGGCGAGCTTGGCGAGC 544
QY 541 GATCGCGCTCGCGCGCGCTTGGCGAGCTTGGCGAGCTTGGCGAGCTTGGCGAGC 600
Db 545 GATCGCGCTCGCGCGCGCTTGGCGAGCTTGGCGAGCTTGGCGAGCTTGGCGAGC 604
QY 601 GCGCGAGTGTGCGAGCTTGGCGAGCTTGGCGAGCTTGGCGAGCTTGGCGAGC 660
Db 605 GCGCGAGTGTGCGAGCTTGGCGAGCTTGGCGAGCTTGGCGAGCTTGGCGAGC 664
QY 661 GCGCTTTCACCGCGCGCTTGGCGAGCTTGGCGAGCTTGGCGAGCTTGGCGAGC 720
Db 665 GCGCTTTCACCGCGCGCTTGGCGAGCTTGGCGAGCTTGGCGAGCTTGGCGAGC 724
QY 721 GAAGATCTACGAGGCGCGCAAGTTTCTGTTTGGCGCGCGCTTGGCGAGCTTGGCGAGC 780
Db 725 GAAGATCTACGAGGCGCGCAAGTTTCTGTTTGGCGCGCGCTTGGCGAGCTTGGCGAGC 784
QY 781 GCGCGCACCATGCGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTG 840
Db 785 GCGCGCACCATGCGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTG 832
QY 841 GTTACGCGACATGTGAAAGCAACGCTTCTTCCAGGTGAGGCGGAGCTTCTTGGCGGA 900
Db 833 GTTACGCGACATGTGAAAGCAACGCTTCTTCCAGGTGAGGCGGAGCTTCTTGGCGGA 892
QY 901 GCAGAGGAGGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 960
Db 893 GCAGAGGAGGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 952
QY 961 CTGGATCCCTTCTTCTTCTGACGGAATCATGCGGCACTCTGTGCTGAGGCTGCCCC 1020
Db 953 CTGGATCCCTTCTTCTTCTGACGGAATCATGCGGCACTCTGTGCTGAGGCTGCCCC 1012
QY 1021 CATCTGAAAGCATATTTCTGTGGCTTGGCTTGGCTTGGCTTGGCTTGGCTTGGCT 1080
Db 1013 CATCTGAAAGCATATTTCTGTGGCTTGGCTTGGCTTGGCTTGGCTTGGCTTGGCT 1072
QY 1081 TTACACAGCTTTTAAACAGAACTTAAACAGAACTTAAACAGAACTTAAACAGAACT 1140
Db 1073 TTACACAGCTTTTAAACAGAACTTAAACAGAACTTAAACAGAACTTAAACAGAACT 1132
QY 1141 ATGAACACAGGG 1152
Db 1133 ATGAACACAGGG 1144

RESULT 4
US-10-333-946-34
; Sequence 34, Application US/10333946
; Publication No. US20040023252A1
; GENERAL INFORMATION:
; APPLICANT: INCYTE GENOMICS, INC.; THORNTON, Michael B.
; APPLICANT: ARVIZU, Chandra S.; LAL, Preeti G.
; APPLICANT: BURFORD, Neil; YUE, Henry
; APPLICANT: GANDHI, Ameena R.; ELLIOTT, Vicki S.

; APPLICANT: RAMKUMAR, Jayalaxmi; BAUGHN, Mariah R.
; APPLICANT: KALLICK, Deborah A.; CHAWLA, Narinder K.
; APPLICANT: HAPILL, April J.A.; YAO, Monique G.
; APPLICANT: LU, Yan; TRIBOULEY, Catherine M.
; APPLICANT: POLICKY, Jennifer L.; KEARNEY, Liam
; APPLICANT: GRAUL, Richard C.; WARREN, Bridget A.
; APPLICANT: LEE, Ernestine A.; DING, Li
; TITLE OF INVENTION: G-PROTEIN COUPLED RECEPTORS
; FILE REFERENCE: PI-0176 USN
; CURRENT APPLICATION NUMBER: US/10/333,946
; CURRENT FILING DATE: 2003-01-22
; PRIOR APPLICATION NUMBER: PCT/US01/23433
; PRIOR FILING DATE: 2001-07-25
; PRIOR APPLICATION NUMBER: US 60/221,478
; PRIOR FILING DATE: 2000-07-27
; PRIOR APPLICATION NUMBER: US 60/223,268
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: US 60/227,054
; PRIOR FILING DATE: 2000-08-21
; PRIOR APPLICATION NUMBER: US 60/231,121
; PRIOR FILING DATE: 2000-09-08
; PRIOR APPLICATION NUMBER: US 60/232,243
; PRIOR FILING DATE: 2000-09-13
; PRIOR APPLICATION NUMBER: US 60/232,691
; PRIOR FILING DATE: 2000-09-15
; PRIOR APPLICATION NUMBER: US 60/235,146
; PRIOR FILING DATE: 2000-09-22
; NUMBER OF SEQ ID NOS: 38
; SOFTWARE: PERL Program
; SEQ ID NO 34
; LENGTH: 1458
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; OTHER INFORMATION: Incyte ID No. US20040023252A1 7476053CB1
US-10-333-946-34

Query Match 92.4%; Score 1064; DB 17; Length 1458;
Best Local Similarity 97.6%; Pred. No. 4.2e-289;
Matches 1121; Conservative 0; Mismatches 0; Indels 27; Gaps 3;

QY 5 ATGAGGCGCTAGAGCTTTCAGTGGCCACCGCGCGGCTTGGCCCTTGGGACCCGAG 64
Db 1 ATGAGGCGCTAGAGCTTTCAGTGGCCACCGCGCGGCTTGGCCCTTGGGACCCGAG 60
QY 65 ACACGAGCGGACCGCGGACCCCAAGCCGAGAGGAGTACTCGTTTCGACCCGCGCGC 124
Db 61 ACACGAG-----CGGACCCCAAGCCGAGAGGAGTACTCGTTTCGACCCGCGCGC 114
QY 125 GCGGTCTCGCGGCGGAGGCGCGGCTTCTCTGTCTTTCACGCTCTGTTGTGAGCGCTG 184
Db 115 GCGGTCTCGCGGCGGAGGCGCGGCTTCTCTGTCTTTCACGCTCTGTTGTGAGCGCTG 174
QY 185 CTAGTGTCTCTGATGCTGCCACTTTCCTGTGGAACCTGCTGTTCCGGTTCACCATCCG 244
Db 175 CTAGTGTCTCTGATGCTGCCACTTTCCTGTGGAACCTGCTGTTCCGGTTCACCATCCG 234
QY 245 CCGGTCCGTGCTTCCACCGCTGCGGATTAATTGTTGCGCTTCGAGCGCGCTCTCGGAC 304
Db 235 CCGGTCCGTGCTTCCACCGCTGCGGATTAATTGTTGCGCTTCGAGCGCGCTCTCGGAC 294
QY 305 GAACTAGTGGCAGCGCTGCGGATGCGACCGAGCTGCGAGTGTGAGTGTGAGTGTGAG 364
Db 295 GAACTAGTGGCAGCGCTGCGGATGCGACCGAGCTGCGAGTGTGAGTGTGAGTGTGAG 354
QY 365 CGTGGCTGCTGGGCGGAGCGCTTGGCCACCGTGTGATCTCTTTCGACCCGCGGAGCTGT 424
Db 355 CGTGGCTGCTGGGCGGAGCGCTTGGCCACCGTGTGATCTCTTTCGACCCGCGGAGCTGT 405
QY 425 CTGTGCTGCGCGCGGCTTGGGAACTGCGGCGGCTTGGGCGGCTTGGGCGGCGGCGG 484
Db 406 CTGTGCTGCGCGCGGCTTGGGAACTGCGGCGGCTTGGGCGGCTTGGGCGGCGGCGG 465

485 ATACACGGGACCTGCGAGCACACGCTGCGCACCCGCGAGCGCGCTCGTTGCTCATGATC 544
Db
466 ATACACGGGACCTGCGAGCACACGCTGCGCACCCGCGAGCGCGCTCGTTGCTCATGATC 525
Qy 545 GCGCTCGCCGGTGGCGCTGCGGCTCATCGCCCTCGCGCGCTGCTCTTTGGCGGGGC 604
Db 526 GCGCTCGCCGGTGGCGCTGCGGCTCATCGCCCTCGCGCGCTGCTCTTTGGCGGGGC 585
Qy 605 GAGGTGCGAGCGCTCGCTCCAGGCTCAGGCTGCGAGCGCGGAAACCTCTATGCGGCC 664
Db 586 GAGGTGCGAGCGCTCGCTCCAGGCTCAGGCTGCGAGCGCGGAAACCTCTATGCGGCC 645
Qy 665 TTCTCCACCGCGCGCTTCCACCTCGCGCTTCCGCTGCGGTGGTGGTGTCTACCGGAAG 724
Db 646 TTCTCCACCGCGCGCTTCCACCTCGCGCTTCCGCTGCGGTGGTGGTGTCTACCGGAAG 705
Qy 725 ATCTACGAGCGCGCAAGTTTCGTTTCGCGCGCGCGCGAGAGCTGCTGCTGCGGTCGCG 784
Db 706 ATCTACGAGCGCGCAAGTTTCGTTTCGCGCGCGCGCGAGAGCTGCTGCTGCGGTCGCG 765
Qy 785 GCCACCATGAGGTGAGGTCCAGGTAAAGGAAGCACTGATGAGGCTGAAGTGGTTC 844
Db 766 GCCACCATGAGGTGAGGTCCAGGTAAAGGAAGCACTGATGAGGCTGAAGTGGTTC 813
Qy 845 ACGGCACATTGCAAGCAAGCGTCTCTCCAGGTGAGCGGGGACTCTCGCGGGAGCAG 904
Db 814 ACGGCACATTGCAAGCAAGCGTCTCTCCAGGTGAGCGGGGACTCTCGCGGGAGCAG 873
Qy 905 AAGGAGGCGGAGCAGCATGATGTTGGGAATCTCGATTGCGGTGTTGTTGCTGCTGG 964
Db 874 AAGGAGGCGGAGCAGCATGATGTTGGGAATCTCGATTGCGGTGTTGTTGCTGCTGG 933
Qy 965 ATCCCTTCTCTGAGCGAACTCATAGCCCACTCTGTGCTCAGGCTGCGCCCAATC 1024
Db 934 ATCCCTTCTCTGAGCGAACTCATAGCCCACTCTGTGCTCAGGCTGCGCCCAATC 993
Qy 1025 TGGAAAGCATATTTCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1084
Db 994 TGGAAAGCATATTTCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1053
Qy 1085 ACAGCTTTTAAAGAACTCAACAATGCTTCAAGAGCGCTCTTTACTAAGCAGATGA 1144
Db 1054 ACAGCTTTTAAAGAACTCAACAATGCTTCAAGAGCGCTCTTTACTAAGCAGATGA 1113
Qy 1145 ACAGAGG 1152
Db 1114 ACAGAGG 1121

RESULT 5
US-10-311-671-25
; Sequence 25, Application US/10311671
; Publication No. US20040072996A1
; GENERAL INFORMATION:
; APPLICANT: INCYTE GENOMICS, INC.
; APPLICANT: LAL, Preeti G.
; APPLICANT: BAUGHN, Mariah R.
; APPLICANT: HAFALIA, April J. A.
; APPLICANT: NGUYEN, Damiel B.
; APPLICANT: GANDHI, Ameena R.
; APPLICANT: KALLICK, Deborah A.
; APPLICANT: GRIFFIN, Jennifer A.
; APPLICANT: YUE, Henry
; APPLICANT: KHAN, Farrah A.
; APPLICANT: ARVIZU, Chandra S.
; APPLICANT: LU, Dyung Aina M.
; APPLICANT: TRIBOULEY, Catherine M.
; APPLICANT: LU, Yan
; APPLICANT: CHAWLA, Narinder K.
; APPLICANT: GRAUL, Richard
; APPLICANT: YAO, Monique G.
; APPLICANT: YANG, Junming

APPLICANT: RAMKUMAR, Javalaxmi
; APPLICANT: AU-YOUNG, Janice K.
; APPLICANT: ELLIOTT, Vicki S.
; APPLICANT: HERNANDEZ, Roberto
; APPLICANT: WALSH, Roderick T.
; APPLICANT: BOROWSKY, Mark L.
; APPLICANT: THORNTON, Michael B.
; APPLICANT: HE, Ann
; TITLE OF INVENTION: G-PROTEIN COUPLED RECEPTORS
; FILE REFERENCE: PI-0131 USN
; CURRENT APPLICATION NUMBER: US/10/311,671
; CURRENT FILING DATE: 2002-12-16
; PRIOR APPLICATION NUMBER: PCT/US01/19275
; PRIOR FILING DATE: 2001-06-15
; PRIOR APPLICATION NUMBER: 60/212,483
; PRIOR FILING DATE: 2000-06-16
; PRIOR APPLICATION NUMBER: 60/213,954
; PRIOR FILING DATE: 2000-06-22
; PRIOR APPLICATION NUMBER: 60/215,209
; PRIOR FILING DATE: 2000-06-29
; PRIOR APPLICATION NUMBER: 60/216,595
; PRIOR FILING DATE: 2000-07-07
; PRIOR APPLICATION NUMBER: 60/218,936
; PRIOR FILING DATE: 2000-07-14
; PRIOR APPLICATION NUMBER: 60/219,154
; PRIOR FILING DATE: 2000-07-19
; PRIOR APPLICATION NUMBER: 60/220,141
; NUMBER OF SEQ ID NOS: 35
; SOFTWARE: PERL Program
; SEQ ID NO 25
; LENGTH: 1119
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; OTHER INFORMATION: Incyte ID No: 7474977CB1
US-10-311-671-25

Query Match 92.0%; Score 1060; DB 12; Length 1119;
Best Local Similarity 97.7%; Pred. No. 5.1e-288;
Matches 1114; Conservative 0; Mismatches 5; Indels 21; Gaps 3;

Qy 5 ATGAGGCGCGCTAGCCTTTTCAGTGGCCACCGCGCGCTTGCCCTTGCCCTGGGACCGCG 64
Db 1 ATGAGGCGCGCTAGCCTTTTCAGTGGCCACCGCGCGCTTGCCCTTGCCCTGGGACCGCG 60
Qy 65 ACCAGCAGCGGACCCCGGACCCCAAGCCCGAGAGGGATCTCGGTTCCGACCCCGAGCGGC 124
Db 61 ACCAGCAG-----CGGACCCCGAGAGGGATCTCGGTTCCGACCCCGAGCGGC 114
Qy 125 GCGTCTCGCGGCGCGAGGCGCGCTTCTCTGTCTTTCACGGTCTCGTGTGAGCGCTG 184
Db 115 GCGTCTCGCGGCGCGAGGCGCGCTTCTCTGTCTTTCACGGTCTCGTGTGAGCGCTG 174
Qy 185 CTAGTGTCTGATCGCTGCCACTTTCTCTGGAACCTGCTGTTCCGGTCAACATCCCG 244
Db 175 CTAGTGTCTGATCGCGCCCACTTTCTCTGGAACCTGCTGTTCCGGTCAACATCCCG 234
Qy 245 CGGTCGCTGCTTCCACCGCGCTGCCCATAACTTGGTGGCCTCGACGGCGCTCTCGGAC 304
Db 235 CGGTCGCTGCTTCCACCGCGCTGCCCATAACTTGGTGGCCTCGACGGCGCTCTCGGAC 294
Qy 305 GAAGTAGTGGCAGCGTGGCGATGCCACCGAGCGCTGGCGAGTGTGTGACGCGGCGA 364
Db 295 GAAGTAGTGGCAGCGTGGCGATGCCACCGAGCGCTGGCGAGTGTGTGACGCGGCGA 354
Qy 365 CGTGGCTGTGGCGGAGCGCTGTGCCACGTGTGATCTCTCTTCGACGCGGAGCGCTGT 424
Db 355 CGTGGCTGTGGCGGAGCGCTGTGCCACGTGTGATCTCTCTTCACG-----TG 405
Qy 425 CTGTGCTGCGCGCGCGCTCGGGAACGTGGCGGCCATCGCCCTGGCGCGCGAGCGGCGC 484

Db 426 CTCATCTTCCCGCGCGAGCGCCCTCTCTGCTTTCACGCTGTTGTGTACTCTA 485
Qy 185 CTAGTGCTGCTGATCGCTGCCACTTTCCTGTGGAACTGCTGTGTCGGTCAACCATCCG 244
Db 486 CTGCTGTTGCTGATCGCTGCCACTTCTTATGGAATCTGTAGTTCTGTGACTATCCGT 545
Qy 245 CGGCTCGTGGCTTCCACCGGTGCGGATCACTTGTGGCTCGACGGCCGCTCTCGGAC 304
Db 546 CGGCTCGGCGCTTCCACCGGTGCGGATCACTTGTGGCTCGACGGCCGCTCTCGGAC 605
Qy 305 GAACTAGTGGCAGCCTGGCGATGCCAGGAGCCTGGCGAGTGGAGCTGTGACCGGGGA 364
Db 606 GTCCCTGGTGGGCTTCTGCTGATGCTCTGAGCCTGTGAGCGAGTTGTCCGCTGGGGA 665
Qy 365 COTCGGCTGCTGGGCGGAGCCTGTGCGAGTGTGATCTCTCTTCCAGCGCGGAGCCTGT 424
Db 666 GTTGGGAGCTAGGACAGAGTCTGTGCCACGTGTGGAATCTCTCTGACG-----TG 716
Qy 425 CTGTGCTCCCGCGCGCTCGGGAACGTGGCGGCATCGCCCTGGGCGCGAGCGGGCC 484
Db 717 TTGTGCTGACCGCGAGCATCTGGAACGTGGCGGCATCGCCCTGGATCGCTACTGGACT 776
Qy 485 ATCAGCGGACCTGGACACACGCTGGGACCGCGAGCGGCGCTGCTGTGCTATGATC 544
Db 777 ATCAGCGGCCCTGGAGTACACGCTGGCGACCGCGAGCGGTGCTTCTGCGCTCATGATC 836
Qy 545 GCGCTCGCGCGGCTGCGGCTCATCGCCCTCGCGCGCTGCTCTTTGGCGGGGCG 604
Db 837 GCGATCAGCTGGGACTGTCGGGCTCATGCTCTCGCCCGCTGCTTTTGGCTGGGCG 896
Qy 605 GAGGTGTGCGAGCTCGGCTCCAGCGTCCAGGTGAGCGGGAAACCTCTATGCCGCC 664
Db 897 GAAGCTATGATGCTCGGCTGACGCTTGCAGGTGAGCGAGGAGCCCTCTATGCTGC 956
Qy 565 TTCTCAGCCCGCGGCTTCCAGCTGCGGCTTGGCGTGGTGGCTGCTTACCGGAG 724
Db 957 TTCTCAGCTCGGAGCCTTCTACCTGCTCTAGCGGTGCTGCTTCTGCTACTGGA 1016
Qy 725 ATCTACGAGGCGGCAAGTTCGTTTCTGGCGCGCGCGGAGCTGTGCTCGCTGCGG 784
Db 1017 ATATACAAAGCGCGAAGTTTCGATTCGTTGCGAGACGGCGGCGGTGTTACCGCTTCC 1076
Qy 785 GCCACCATGAGGTGAGTCCAAAGTAAAGAACACCTGATGAGGTGAAGTGTGTTTC 844
Db 1077 GCCACCAAGC-----AGGCAAGGAAGCACCTCGGAGTCTGAGATGTTGTTTC 1124
Qy 845 ACGCATATGCAAGCAACGCTGCTCTTCCAGGTGAGCGGGAGTCTTGGCGGAGAG 904
Db 1125 ACAGCCGTCGCGAGCAACAGTACCTTCCAGAACAGCGGAGACTCTCTGGCGGAGCAG 1184
Qy 905 AAGGAGGCGGAGCAGCATGATGTTGGGAATCTGATTGGCGGTTTGTGCTGTGCTGG 964
Db 1185 AAGGAGGCGGAGCAGCATGATGTTGGGATCTTGAATGGCGGTTTGTGCTGTGTTGG 1244
Qy 965 ATCCCTTCTTCTGAGCGGAACATCAGGCCATCTGTGCTGAGCTGCCCCCCTATC 1024
Db 1245 ATCCCTTCTTCTGAGCGGAGTCTCAGCCGCTCTGTGCTGAGCTGCCCCCCTATC 1304
Qy 1025 TGGAAAGCATATTTCTGCTGGCTGGCTACTCCTCAATCTTCTTCAACCCCTGATTTAC 1084
Db 1305 TGGAAAGCATATTTCTGCTGGCTGGATATTTCAATCTTCTTCAACCCCTGATTTAC 1364
Qy 1085 ACAGCTTTTAAAGAACTCAACAAATGCTTCAAGAGCCTCTTACTAAGCAGATGA 1144
Db 1365 ACTGCTTTTAAAGAACTCAACAAATGCTTCAAGAGCCTCTTACTAAGCAGATGA 1424

RESULT 10

US-10-225-567A-445

; Sequence 445, Application US/10225567A

; Publication No. US20030113798A1

; GENERAL INFORMATION:

; APPLICANT: LifeSpan Biosciences

; APPLICANT: Brown, Joseph P.
; APPLICANT: Burmer, Glenna C.
; APPLICANT: Roush, Christine L.
; TITLE OF INVENTION: ANTIGENIC PEPTIDES AND ANTIBODIES FOR G PROTEIN-COUPLED RECEPTORS
; FILE REFERENCE: 1920-4-4
; CURRENT APPLICATION NUMBER: US/10/225,567A
; CURRENT FILING DATE: 2001-12-19
; PRIOR APPLICATION NUMBER: 60/257,144
; PRIOR FILING DATE: 2000-12-19
; NUMBER OF SEQ ID NOS: 2292
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 445
; LENGTH: 1074
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-225-567A-445

Query Match 43.2%; Score 497.6; DB 15; Length 1074;
Best Local Similarity 70.6%; Pred. No. 1.le-129;
Matches 729; Conservative 0; Mismatches 279; Indels 24; Gaps 4;

Qy 113 ACCCGAGCGCGCTCTCGCGGCGAGGGCGCCCTTCTGTGTTCACGGTCTG 172
Db 67 AGCCTCGGCAAGACGACCTGCGCCCGAGCTCGGCCCTGCTCTCGGTCTCGGAGTGGT 126
Qy 173 GTGTGAGCGCTAGTGTCTGCTGATCGCTGCCACTTTCCTGTGGAACCTGTGTTCCG 232
Db 127 ATTCTACCTTGTGGGCTTTCTGTGGCGGCGAGTTCGCTGGAACCTGTGTTGCTG 186
Qy 233 GTACCATTCGCGGGTCCGTGCTTCCACCGCTGCGGCAATACTTGGTGGCTCGACG 292
Db 187 GCGACCATCTCCGTGTACGACACCTTCCACGCGTGGCCCAACACCTGTGGCATTCATG 246
Qy 293 GCGCTTCGAGCAAACTAGTGGCAGCGCTGGCGATGCCACGAGCCCTGGGAGTGAAGT 352
Db 247 GCGCTCTGAGTGTCTGTGGCGCGCTGTCTATGCGCTGAGCCCTGGTGCATGAGCTG 306
Qy 353 TCGACCGGCGAGCTCGGCTGCTGGCGGAGCCTGTGCCACGTGTGGATCTCTTCGAC 412
Db 307 T---CCGGCGCCGCTGGCAGTAGGTGGAGGTGTGCCAGCTTGGATCCGCTGGAC 363
Qy 413 GCCGAGCGCTGTGTGCTGCCCGCGGCTCGGGAAGTGGCGGCATCGCCCTGGGCG 472
Db 364 G-----TGCTTGTGCACGCCAGCATCTGGAACGTGACGGCATAGCCCTGGAC 414
Qy 473 CGCACCGGCGCATCACAGCGACCTCGAGCAACAGCTGGCGCACCCGAGCGCGCTCG 532
Db 415 CGCTACTGTGTCATCAGCGGCCACATGGAAATACACGCTCGGCAACCCGCAAGTGGCTCC 474
Qy 533 TTGCTCATGATCGGCTCGCCGCGGTGCGCTCGGCGGCTCATCGCCCTCGCGCGCTGCTC 592
Db 475 AACGTATGATCGGCGCTACCTGGGCACTCTCCGCTGTCTCTCTGCGCCCGCTGCTT 534
Qy 593 TTTGGCGGCGGAGGTGTGCGAGCTCGGCTCCAGGCTGCCAGGTGAGCGCGGAACCC 652
Db 535 TTTGGCTGGGAGAGAGTACTCTGAGGCGAGCGAGAGTGCAGGTAAAGCGGAGCCT 594
Qy 653 TCCTATCGCGCTTCTCCACCGCGCGCTTCCACTGCGCGCTTGGCGGTGGTGGTCCGCTT 712
Db 595 TCCTACCGCGTGTCTCCACCGTAGGCGCTTCTACCTGCGCTCTGTGTGGTCTCTTC 654
Qy 713 GTCTACCGGAGATCTACGAGGCGCAAGTTTGGTTTCGCGCCCGCGCGAGAGCTGTG 772
Db 655 GTGTACTGGAAGATCTACAAGGCTGCCAAGTTCGCGTGGGCTCCAGGAAGACCAATAGC 714
Qy 773 CTGCGGTTCGCGGCACCATGCAAGTGTGAGTCTCAAGGTAAAGGAAGCACCTGTAGGCT 832
Db 715 GT-----CTCACCATATCGAAGCTGTGGAGTGAAGACTCTGCCAACAGCCCC 765
Qy 833 GAAGTGTGTTCAGGCAATTCGAAAGCAACGGTGTCTTCCAGGTGAGCGGAGCTCC 892
Db 766 CAGATGTTGTTACGGTCC---GCCAGCCACCGTCACTTCCAGCCAGGAAGGAGCAAG 822


```
; PRIOR FILING DATE: 2002-06-24
; PRIOR APPLICATION NUMBER: US 60/392,480
; PRIOR FILING DATE: 2002-06-28
; PRIOR APPLICATION NUMBER: US 60/394,128
; PRIOR FILING DATE: 2002-07-03
; PRIOR APPLICATION NUMBER: US 60/399,783
; PRIOR FILING DATE: 2002-07-31
; PRIOR APPLICATION NUMBER: US 60/403,221
; PRIOR FILING DATE: 2002-08-13
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 62
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 15
; LENGTH: 1159
; TYPE: DNA
; ORGANISM: Homo Sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (64)...(1137)
US-10-352-684A-15

Query Match      43.2%; Score 497.6; DB 16; Length 1159;
Best Local Similarity 70.6%; Pred. No. 1.1e-129;
Matches 729; Conservative 0; Mismatches 279; Indels 24; Gaps 4;

QY 113 ACCCGAGCGCGCGCTGCTGCGGCGCGAGGCGCGCCCTTCTCTCTTCAAGTCTCG 172
Db 130 AGCCTCGGAAGACAGACTGCGCCAGCTGCGCCCTGCTCTCGTCTTCGGAGTCTT 189
QY 173 GTGGTGAAGCTGCTAGTGTCTGTGATCGTGTGCACTTTCTGTGGAACCTGCTGTTCCG 232
Db 190 ATTCTCACCTTGTGGGCTTTCTGTGGCGGAGCTTTCGCTGGAACCTGCTGTGTGCTG 249
QY 233 GTACCATCCGCGGCTGCTGCTTCCACCGGTCGCGCATTAACCTTGTGCTGCTGAG 292
Db 250 GCGACCATCTCGGTGTACGACCTTCCACCGGTCGCGCATTAACCTTGTGCTGCTGAG 309
QY 293 GCGCTCTCGGACCACTAGTGGCAGCGCTGGCGATGCCACCGAGCTGGCGAGTGAAGCTG 352
Db 310 GCGCTCTCGAGTCTGCTGGTGGCGGCTGGTCTGATGCGCTGAGCTTGTGCTGAGCTG 369
QY 353 TCGACCGGCGAGTCTGCTGCTGGCGGAGCTGTGCGACGCTGTGCGATCTCTCTTCGAC 412
Db 370 T---CGGCGCGCGCTGGCAGCTAGTTCGAGGCTGTGCGAGCTTGTGATTCGCGTGGAC 426
QY 413 GCGGAGCTGTGTGTGCTGCCCCCGGCGCTCGGGAAGCTGCGGCGCATGCGCCCTGGG 472
Db 427 G-----TGCTTTTGCTGACGGCGCAGCATCTGGAACGTGACGCGCATAGCCCTGGAC 477
QY 473 CGCGAGCGGCGCATCACAGGCACTTGCGACACACGCTGCGACCGCGAGCGCGCTCG 532
Db 478 CGTACTGTGTCCATCAGCGGCACATGGAATACACGCTCGCACCGCGAAGTSGTCTCC 537
QY 533 TTGCTCATGATCGCTGCGCGCGGCTGCGTGTGCGGCTATGCGCCCTGCGCGGCTGCTC 592
Db 538 AAGCTCATGATCGCTGCTGCGGCTGCTGCGGCTGCTGCTGCTGCTGCTGCTGCTGCT 597
QY 593 TTGCGCGGCGGAGTGTGCGAGCTCGCTCCAGGCTGCGGAGTGAAGTGAAGCGGAGCC 652
Db 598 TTGCGTGGGAGAGACGCTACTCTGAGGCGAGGAGTGCAGGTAAGCGCGGAGCCT 657
QY 653 TCCTATGCGCGCTTCTCCACCGCGCGCTTCCACCTGCGGCTTGGCGTGTGCGGCTTT 712
Db 658 TCCTAGCGCGTGTCTCCACCGTAGGCGCTTCTACTGCGCTCTGTGTGCTGCTCTTC 717
QY 713 GTCTACCGGAAGATCTACGAGGCGGCAAGTTTGTGTTTGGCGCGCGCGGAGAGCTGTG 772
Db 718 GTCTACTGGAAGATCTACAGAGGCTGCAAGTTTCCGGTGGGCTTCCAGGAAGCAATAGC 777
QY 773 CTGCGCGTTCGCGCACCATGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTG 832
Db 778 GT-----CTACCCCATATCCGAAGCTGTGAGGTGAAGACTCTGCAAAACAGCCC 828
```

```
QY 833 GAAGTGGTGTTCACGGCACATTTGCAAAAGCAACGGTGTCTTCCAGGTGAGCGGGGACTCC 892
Db 829 CAGATGGTGTTCACGGTCC---GCCACGCGCACCGTTCACCTTCCAGCCAGAGGGGACAG 885
QY 893 TGGCGGGAGCAGAGAGAGAGCGAGCAGCCATGATGGTGGGAATTTCTGATTGCGCGTGT 952
Db 886 TGGCGGGAGCAGAGAGAGAGAGCGGCGCGCCCTCATGGTGGGCACTCTCATTTGGCGTGT 945
QY 953 GTGCTGTGCTGGATCCCTTCTTCTGAGCGGAACCTCATCAGCCCACTCTGTGCTGCGAC 1012
Db 946 GTGCTGTGCTGGATCCCTTCTTCTCACCAGACTCATCAGTCCCTCTGCTGCTGTGAC 1005
QY 1013 CTGCCCCCATCTGGAAGCATATTTCTGTGGCTTGGCTACTCAATTTCTTCTTCAAC 1072
Db 1006 ATCCCGGCGCATCTGGAAGCATCTTCTGTGGTGGCTACTCAAACTCTTCTTTTAA 1065
QY 1073 CCCCTGATTTACACAGCTTTTAAAGAACTCAACAATGCTTCAAGAGCTCTTTACT 1132
Db 1066 CCCCTGATCTATACGCGCTTTCAAGAAGAACTCAACAAGCGCTTCAAGAACTTCTTTCT 1125
QY 1133 AAGCAGAGATGA 1144
Db 1126 AGGCAACACTGA 1137

RESULT 14
US-09-823-245A-617
; Sequence 617, Application US/09823245A
; Publication No. US20020039760A1
; GENERAL INFORMATION:
; APPLICANT: Wong, Gordon G.
; APPLICANT: Clark, Hilary
; APPLICANT: Fechtel, Kim
; APPLICANT: Agostino, Michael J.
; APPLICANT: Howes, Steven H.
; APPLICANT: Resnick, Richard J.
; APPLICANT: Gulukota, Kamalakr
; APPLICANT: Graham, James R.
; APPLICANT: Genetics Institute, Inc.
; TITLE OF INVENTION: POLYNUCLEOTIDES ENCODING NOVEL SECRETED PROTEINS
; FILE REFERENCE: GIN 6401
; CURRENT APPLICATION NUMBER: US/09/823,245A
; CURRENT FILING DATE: 2001-03-29
; PRIOR APPLICATION NUMBER: 60/194,941
; PRIOR FILING DATE: 2000-04-06
; NUMBER OF SEQ ID NOS: 631
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 617
; LENGTH: 2700
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-823-245A-617

Query Match      43.1%; Score 496; DB 13; Length 2700;
Best Local Similarity 70.5%; Pred. No. 4.1e-129;
Matches 728; Conservative 0; Mismatches 280; Indels 24; Gaps 4;

QY 113 ACCCGAGCGCGCGCTGCTGCGGCGCGAGGCGCGCCCTTCTCTCTTCAAGTCTCG 172
Db 371 AGCCTCGGAAGACAGCACTGCGCCAGCTCGCCCTGCTCTCGGCTTTCGGAGTCTT 430
QY 173 GTGGTGAAGCTGTAGTGTGCTGATCGTGTGCACTTTTCTGTGGAACCTGCTGTTCCG 232
Db 431 ATTCTCACCTTGTGGGCTTTCTGTGGCGGCGACGTTGCGCTGGAACCTGCTGTGCTG 490
QY 233 GTACCATCCGCGGCTGCTGCTTCCACCGGTCGCGCATTAACCTTGTGCGCTCGAG 292
Db 491 GCGACCATCTTCTGCTGATGACACTTTCCACCGGTCGCGCGTCCCAACCTGTTGGCAT 550
QY 293 GCGCTCTCGGACGAACTAGTGGCAGCGCTGGCGATGCCACCGAGCTTGGCGAGTGTG 352
Db 551 GCGCTCTCGGATCTCTGTTGGTGGCGCGCTGCTCATGCCGTGAGCCTTGTGTCAGAGCTG 610
```

QY	353	TCGACGGGCGACGTGCGTGTCTGGCCCGGACCGCTGTGCACAGTGTGGATCTCTCTTCGAC	411
Db	611	T----CCGGCCGCGCTGGCAGCTAGGTCCGAGGCTGTGCCAGCTTTGGATCGCGTGCAC	667
QY	413	GCCGAGCGTGTCTGTGTGCTGCCCGCCGCGCTCGGGACGTGGCGGCCATCGCCCTGGCC	472
Db	668	G-----TGCTTTGTGCACGGCCAGCATCTGGAACGTGAAGCCATATAGCCCTGGAC	718
QY	473	GCGACGGGGCCATCACACGGCACCTGCAGCACACGCTGCGCACCGCAGCGCGGCCCTCG	532
Db	719	CGCTACTGGTCCATCAACGCGCCACATGAAATACACGCTCCGCACCCGCAAGTGGCTCTCC	778
QY	533	TTGCTCATGATCGCGCTCGCCCGGCTGGCTGGCGCTCATCGCCCTCGCGCGCGTGCTC	592
Db	779	AACGTCATGATGGCGCTCACCTGGGCATCTCCGCTGTCACTCTCTGGCCCGCGTGCTT	838
QY	593	TTTGGCCGGGCGAGTGTGCAGCCTCGGCTCCAGCGCTGCCAGGTGCCAGGTGAGCGGGACCC	652
Db	839	TTTGGCTGGGAGAGACGTACTCTGAGGCGAGCGAGGAGTGCCAGGTAAAGCCCGAGGCT	898
QY	653	TCCTATGCGCGCTTCTCCACCGCGGCGCTTCCACTGCGCGCTTGGCGTGGTCCCGTTT	712
Db	899	TCCTACGCGGTGTCTCCACCGTAGGCGCCTTCTACCTGCGCTCTGTGTGCTGCTCTTC	958
QY	713	GTCCTACCGGAAGACTACGAGCGCGCAAGTTTCGTTTCGGCCGCGCGCGAGAGCTGTG	772
Db	959	GTGTACTGGAAGATCTACAAGGCTGCCAAGTTCGCGGTGGGCTCCAGGAAGACCAATAGC	1018
QY	773	CTGCGCTTGC CGGCCACCATGCAGGTGAGGTCCAAAGGTAAAGGAAGCACCTGATGAGGCT	832
Db	1019	GT-----CTACCCATATCCGAGCTGTGAGGTGAAGGACTCTGCCAAACAGCCC	1069
QY	833	GAAGTGTGTTTCAGGCACATGTCAAAGCAACGGTGTCTTCTCCAGGTGAGCGGGGACTCC	892
Db	1070	CAGATGTGTTCACGGTCC---GCCAGCCACCGCTCACTTCCAGCCAGAAGGSGGACACG	1126
QY	893	TGGCGGAGCAGAGGAGAGCGCAGCAGCCATGATGTTGGGAATCTGATTGGCGTGTTT	952
Db	1127	TGGCGGAGCAGAAGGAGCAGCGCGCCGCCCTCATGTGGGCAFCCTCATTTGGCGTGTTC	1186
QY	953	GTGCTGTCTGGATCCCGCTTCTTCCTGACGGAACTCATACGCCCACTCTGTGCTCGCAGC	1012
Db	1187	GTGCTGTCTGGATCCCGCTTCTTCTCACCGAGCTCATCAGTCCCTCTGCTCTCTGTGAC	1246
QY	1013	CTGCCCCCATCTGGAAGAAGCATATTCTGTGGCTTGGCTACTCCAAATCTTTCTTCAAC	1072
Db	1247	ATCCCCGCCATCTGGAAGAAGCATCTCTGTGGCTTGGCTACTCCAACTCCTTCTTTAAC	1306
QY	1073	CCCTGATTTTACACAGCTTTTAAACAAGACTACAAACAATGCCCTTCAGAGGCTCTTTACT	1132
Db	1307	CCCTGTATCTATACGGCTTTCAACAAGACTACAAAGCGCCCTTCAGAACTCTTTTCT	1366
QY	1133	AAGCAGAGATGA	1144
Db	1367	AGGCAACACTGA	1378

RESULT 15

US-09-826-509-446
; Sequence 446, Application US/09826509
; Publication No. US20030204073A1
; GENERAL INFORMATION:
; APPLICANT: Lehmann-Bruinsma, Karin
; APPLICANT: Liaw, Chen W.
; APPLICANT: Lin, I-Lin
; TITLE OF INVENTION: No. US20030204073A1-Endogenous, Constitutively Activated Known G
; TITLE OF INVENTION: Protein-Coupled Receptors
; FILE REFERENCE: ASEN-207
; CURRENT APPLICATION NUMBER: US/09/826,509
; CURRENT FILING DATE: 2001-04-05
; PRIOR APPLICATION NUMBER: 60/195,747
; PRIOR FILING DATE: 2000-04-07
; PRIOR APPLICATION NUMBER: 09/170,496

Wed Sep 8 09:18:16 2004

QY	1013	CTGCCCCCATCTGGAAGCATATTCTGTGGCTTGGCTACTCCAAATTCITTTCTTCAAC	1072
Db	943	ATCCCCGGCATCTGGAAGCATCTCTGTGGCTTGGCTACTCCAACTCCTTCTTTAAC	1002
QY	1073	CCCCTGATTTACACAGCTTTTAAACAAGAACTACAACAATGCCTTCAAGAGCCTCTTTACT	1132
Db	1003	CCCCTGATCTATACGGCTTTCAACAAGAACTACAACAAGCCTTCAAGAACTTCTTTCT	1062
QY	1133	AAGCAGAGATGA	1144
Db	1063	AGGCAACACTGA	1074

Search completed: September 7, 2004, 18:33:10
Job time : 592 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 7, 2004, 14:14:17 ; Search time 4454 Seconds
(without alignments)
11210.412 Million cell updates/sec

Title: US-09-976-782-15
Perfect score: 1152
Sequence: 1 cgcctgagggccgctagcc.....aagcagagatgaacacaggg 1152

Scoring table: IDENTITY.NUC
Gapop 10.0 , Gapext 1.0

Searched: 3470272 seqs, 21671516995 residues
Total number of hits satisfying chosen parameters: 6940544

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : GenEmbl.*

- 1: gb.ba.*
- 2: gb.htg.*
- 3: gb.in.*
- 4: gb.om.*
- 5: gb.ov.*
- 6: gb.pat.*
- 7: gb.ph.*
- 8: gb.pl.*
- 9: gb.pr.*
- 10: gb.ro.*
- 11: gb.sts.*
- 12: gb.sy.*
- 13: gb.un.*
- 14: gb.vi.*
- 15: em.ba.*
- 16: em.fun.*
- 17: em.hum.*
- 18: em.in.*
- 19: em.mu.*
- 20: em.om.*
- 21: em.or.*
- 22: em.ov.*
- 23: em.pat.*
- 24: em.ph.*
- 25: em.pl.*
- 26: em.ro.*
- 27: em.sts.*
- 28: em.un.*
- 29: em.vi.*
- 30: em.htg.hum.*
- 31: em.htg.inv.*
- 32: em.htg.other.*
- 33: em.htg.mus.*
- 34: em.htg.pln.*
- 35: em.htg.rod.*
- 36: em.htg.mam.*
- 37: em.htg.vrt.*
- 38: em.sy.*
- 39: em.htgo.hum.*
- 40: em.htgo.mus.*
- 41: em.htgo.other.*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB	ID	Description
1	1152	100.0	1152	6	AX527761	AX527761 Sequence
2	1068	92.7	1150	6	AX301902	AX301902 Sequence
3	1064.8	92.4	1150	6	AX301904	AX301904 Sequence
4	1064	92.4	1458	6	AX375244	AX375244 Sequence
5	1060	92.0	1119	6	AX342670	AX342670 Sequence
6	1016	88.2	1155	6	AX526745	AX526745 Sequence
7	756.2	65.6	1086	6	AX646417	AX646417 Sequence
8	756.2	65.6	44312	9	AB065453	AB065453 Homo sapi
9	749.8	65.1	112883	9	AC009404	AC009404 Homo sapi
10	749.8	65.1	184328	2	AC011638	AC011638 Homo sapi
11	745	64.7	977	9	HSA308679	AJ308679 Homo sapi
12	732.8	63.6	2036	6	A39678	A39678 Sequence 1
13	732.8	63.6	2061	10	MM5HT5BR	X69867 M.musculus
14	729.8	63.4	2226	6	AR080627	AR080627 Sequence
15	729.8	63.4	2240	10	RAT5HT5B	L10073 Rattus norv
16	497.6	43.2	1074	6	AX549160	AX549160 Sequence
17	496	43.1	1074	9	AF498985	AF498985 Homo sapi
18	494.4	42.9	2911	9	AK094255	AK094255 Homo sapi
19	492.8	42.8	1074	6	AX280823	AX280823 Sequence
20	488.2	42.4	207556	10	AC128353	AC128353 Rattus no
21	487.8	42.3	240622	2	AC073793	AC073793 Mus muscu
22	484.4	42.0	1073	6	A74295	A74295 Sequence 7
23	484.4	42.0	1073	6	AR038836	AR038836 Sequence
24	424.6	36.9	1686	6	A74289	A74289 Sequence 1
25	424.6	36.9	1686	6	AR038830	AR038830 Sequence
26	424.6	36.9	1686	6	AX392847	AX392847 Sequence
27	424.6	36.9	1686	10	MM5HT5SE	Z18278 Mus musculu
28	416	36.1	1954	10	RAT5HT5A	L10072 Rattus norv
29	416	36.1	2155	6	AR080628	AR080628 Sequence
30	394.8	34.3	110000	2	AC073744_0	AC073744 Mus muscu
31	378.8	32.9	203822	2	AC146825	AC146825 Otolemur
32	378.2	32.8	636	6	AX188694	AX188694 Sequence
33	377.2	32.7	194024	2	AC146728	AC146728 Otolemur
34	350.4	30.4	184328	2	AC011638	AC011638 Homo sapi
35	348.8	30.3	459	9	HSA308680	AJ308680 Homo sapi
36	334.4	29.0	336	6	A39680	A39680 Sequence 3
37	320	27.8	840	9	HSSHT5A1	X81411 H.sapiens D
38	320	27.8	840	11	G28709	G28709 sWSS3980 Er
39	318.4	27.6	148832	2	AC074285	AC074285 Homo sapi
40	318.4	27.6	159041	9	AC093726	AC093726 Homo sapi
41	316.8	27.5	207644	2	AC146072	AC146072 Pan trogl
42	316.4	27.5	310072	2	AC111733	AC111733 Rattus no
43	297	25.8	297	6	AX188693	AX188693 Sequence
44	273.6	23.8	110000	2	AC111011_2	Continuation (3 of
45	273.6	23.8	205677	2	AC068651	AC068651 Mus muscu

ALIGNMENTS

RESULT 1
AX527761
LOCUS
DEFINITION Sequence 15 from Patent WO0230974.
ACCESSION AX527761
VERSION AX527761.1 GI:25172284
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE
AUTHORS
1
Grosse, W.M., Alsobrook, J.P., Lepley, D.M., Burgess, C.E., Mishra, V.,
Kekuda, R., Li, L., Padigaru, M., Shinkets, R.A., Zerkhus, B.D.,
Spytek, K.A., Edinger, S., Gerlach, V., Macdougall, J., Stone, D.,

Gunther, E. and Ellerman, K.
Proteins and nucleic acids encoding same
Patent: WO 0230974-A 15 18-APR-2002;
Curagen Corporation (US)

FEATURES
Location/Qualifiers

1..1152

/organism="Homo sapiens"

/mol_type="unassigned DNA"

/db_xref="taxon:9606"

ORIGIN

Query Match 100.0%; Score 1152; DB 6; Length 1152;
Best Local Similarity 100.0%; Pred. No. 2.7e-182;
Matches 1152; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CGCCATGAGGCGCTAGCTTTAGTGGCCACCGCGCGCTTGCCTTGCCTGGGACC 60
DB 1 CGCCATGAGGCGCTAGCTTTAGTGGCCACCGCGCGCTTGCCTTGCCTGGGACC 60

QY 61 CGAGACACGACGACGACGACGACGACGACGACGACGACGACGACGACGACGAC 120
DB 61 CGAGACACGACGACGACGACGACGACGACGACGACGACGACGACGACGACGAC 120

QY 121 CGCGCGCTCTGCGCGCGCGCGCGCGCGCTTCTCTCTCTCTCTCTCTCTCTCTCT 180
DB 121 CGCGCGCTCTGCGCGCGCGCGCGCGCGCTTCTCTCTCTCTCTCTCTCTCTCTCT 180

QY 181 GCTGCTAGTGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 240
DB 181 GCTGCTAGTGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 240

QY 241 CGCGCGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 300
DB 241 CGCGCGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 300

QY 301 GGACGAACCTAGTGGACGCTGCGATGCCACCGACCTGGGAGTGTGCTGCTGCTGCT 360
DB 301 GGACGAACCTAGTGGACGCTGCGATGCCACCGACCTGGGAGTGTGCTGCTGCTGCT 360

QY 361 GCGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 420
DB 361 GCGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 420

QY 421 CTGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 480
DB 421 CTGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 480

QY 481 GGCCATCACACGACCTGTCAGCACACGCTGCGCACCCGACCGCGCGCTGCTGCTGCT 540
DB 481 GGCCATCACACGACCTGTCAGCACACGCTGCGCACCCGACCGCGCGCTGCTGCTGCT 540

QY 541 GATCGGCTCGCGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 600
DB 541 GATCGGCTCGCGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 600

QY 601 GCGCGAGTGTGCGACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 660
DB 601 GCGCGAGTGTGCGACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 660

QY 661 CGCCTTCTCCACCGCGCGCTTCCACCTGCGCTGCGCTGCGCTGCGCTGCGCTGCGCT 720
DB 661 CGCCTTCTCCACCGCGCGCTTCCACCTGCGCTGCGCTGCGCTGCGCTGCGCTGCGCT 720

QY 721 GAGATCTACGAGCGCGCAAGTTTCTGTTTGGCGCGCGCGCGCGCGCGCTGCTGCTGCT 780
DB 721 GAGATCTACGAGCGCGCAAGTTTCTGTTTGGCGCGCGCGCGCGCGCGCTGCTGCTGCT 780

QY 781 GCGCGCACCATGACAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAG 840
DB 781 GCGCGCACCATGACAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAG 840

QY 841 GTTCAGGACATGTCAGAACACGCTGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 900

DB 841 GTTCACGACATTCGAAAGCAACGGTGTCTCTCCAGGTGACGGGGGATCTCTGGGGGA 900
QY 901 GCAGAGGAGAGGCGAGCAGCATCATGTGGGAATTTCTGATTGGCGTGTGTTGCTGTG 960
DB 901 GCAGAGGAGAGGCGAGCAGCATCATGTGGGAATTTCTGATTGGCGTGTGTTGCTGTG 960
QY 961 CTGGATCCCTTTCTCTCTGAGCGAACTCATCAGCCCACTCTGTGCTGAGCCTGCCCC 1020
DB 961 CTGGATCCCTTTCTCTCTGAGCGAACTCATCAGCCCACTCTGTGCTGAGCCTGCCCC 1020

QY 1021 CATCTGGAAGACATATTTCTGTGGTGGTCTCTCAATCTTCTTCAACCCCTGAT 1080
DB 1021 CATCTGGAAGACATATTTCTGTGGTGGTCTCTCAATCTTCTTCAACCCCTGAT 1080

QY 1081 TTACACAGCTTTTAAACAAGAACTACAACAATGCTTCAAGAGCCTCTTTACTAAGCAG 1140
DB 1081 TTACACAGCTTTTAAACAAGAACTACAACAATGCTTCAAGAGCCTCTTTACTAAGCAG 1140

QY 1141 ATGAACACAGGG 1152
DB 1141 ATGAACACAGGG 1152

RESULT 2

AX301902 1150 bp DNA linear PAT 30-NOV-2001
DEFINITION Sequence 9 from Patent WO0174851.
ACCESSION AX301902
VERSION AX301902.1 GI:17382952
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE
AUTHORS Majumder, K., Spaderna, S.K., Taupier, R.J., Padigaru, M.,
Burgess, C.B., Shmukets, R.A., Spytek, K.A., Liu, X., Fatturajan, M. and
Gusev, V.Y.
TITLE Novel proteins and nucleic acids encoding same
JOURNAL Patent: WO 0174851-A 9 11-OCT-2001;
Curagen Corporation (US)

FEATURES
source 1..1150
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
24..1136
/note="unassigned protein product"
/codon_start=1
/protein_id="CAD13087.1"
/db_xref="GI:17382953"
/db_xref="REMTREMBL:CAD13087"

CDS

1..1136
/note="unassigned protein product"
/codon_start=1
/protein_id="CAD13087.1"
/db_xref="GI:17382953"
/db_xref="REMTREMBL:CAD13087"

ORIGIN

Query Match 92.7%; Score 1068; DB 6; Length 1150;
Best Local Similarity 97.7%; Pred. No. 2.7e-168;
Matches 1125; Conservative 0; Mismatches 27; Indels 27; Gaps 3;

QY 1 CGCCATGAGGCGCTAGCTTTAGTGGCCACCGCGCTTGCCTTGCCTGGGACC 60
DB 20 CGCCATGAGGCGCTAGCTTTAGTGGCCACCGCGCTTGCCTTGCCTGGGACC 79

QY 61 CGAGACACGACGACGACGACGACGACGACGACGACGACGACGACGACGACGAC 120
DB 80 CGAGACACGACGACGACGACGACGACGACGACGACGACGACGACGACGACGAC 133


```
/db_xref="taxon:9606"
/note="Incyte ID No: 7476053CB1"

Query Match 92.4%; Score 1064; DB 6; Length 1458;
Best Local Similarity 97.6%; Pred. No. 1.2e-167;
Matches 1121; Conservative 0; Mismatches 0; Indels 27; Gaps 3;

QY 481 GGCCATCACAGCGACCTGACAGCACACGCTGCGCACCGCCAGCGCGCTGCTGCTCAT 540
Db 485 GGCCATCACAGCGACCTGACAGCACACGCTGCGCACCGCCAGCGCGCTGCTGCTCAT 544
QY 541 GATCGCGCTCGCCCGGGTGGCGTGGCGCTCATCGCCCTCGCGCGCTGCTGCTGCGCG 600
Db 545 GATCGCGCTCGCCCGGGTGGCGTGGCGCTCATCGCCCTCGCGCGCTGCTGCTGCGCG 604
QY 601 GGGCGAGGTGCGAGCGCTCGGCTCAGGCTGCGAGCTGCGAGTGGCCGGACCTCTATGC 660
Db 605 GGGCGAGGTGCGAGCGCTCGGCTCAGGCTGCGAGTGGCCGGACCTCTATGC 664
QY 661 GCGCTTCTCCACCCCGCGCGCTTCCACCTGCGGTTGGCGTGGTGGTGGTGGTGGT 720
Db 665 GCGCTTCTCCACCCCGCGCGCTTCCACCTGCGGTTGGCGTGGTGGTGGTGGTGGT 724
QY 721 GAAGATCTACGAGGGCGGCAAGTTTCTGTTTGGCGCGCGCGCGAGCTGCTGCCGTT 780
Db 725 GAAGATCTACGAGGGCGGCAAGTTTCTGTTTGGCGCGCGCGAGCTGCTGCCGTT 784
QY 781 GCGGCGACCATGAGGTGAGTCCAGTAAAGGACACCTGATGAGGCTGAGTGGT 840
Db 785 GCGGCGACCATGAGTCCAGTAAAGGACACCTGATGAGGCTGAGTGGT 840
QY 841 GTTTCAGCGCACATTGCAAGCAACCGTGTCTTCCAGGTGAGCGGGACTCCTGCGGGA 900
Db 833 GTTTCAGCGCACATTGCAAGCAACCGTGTCTTCCAGGTGAGCGGGACTCCTGCGGGA 892
QY 901 CGAAGAGGAGCGGAGCGAGCATGATGTTGGAAATTCGATGCGGTTGTTGTTGTTG 960
Db 893 CGAAGAGGAGCGGAGCGAGCATGATGTTGGAAATTCGATGCGGTTGTTGTTGTTG 952
QY 961 CTGGATCCCTTCTTCTGACGAACTCATGACCCACTCTGCTGCGAGCTGCGCCC 1020
Db 953 CTGGATCCCTTCTTCTGACGAACTCATGACCCACTCTGCTGCGAGCTGCGCCC 1012
QY 1021 CATCTGGAAGACATATTTCTGCTGCTTGGCTTACTTCAATTTCTTCAACCCCTGAT 1080
Db 1013 CATCTGGAAGACATATTTCTGCTGCTTGGCTTACTTCAATTTCTTCAACCCCTGAT 1072
QY 1081 TTACACAGCTTTTAAACAGAACTTACAAATGCTTCAAGAGCTTTTAAAGCAGAG 1140
Db 1073 TTACACAGCTTTTAAACAGAACTTACAAATGCTTCAAGAGCTTTTAAAGCAGAG 1132
QY 1141 ATGAACACAGGG 1152
Db 1133 ATGAACACAGGG 1144

RESULT 4
AX375244
LOCUS
DEFINITION
Sequence 34 from Patent WO0210387.
ACCESSION
AX375244
VERSION
AX375244.1 GI:19169962
KEYWORDS
Homo sapiens (human)
SOURCE
Homo sapiens
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1
REFERENCE
AUTHORS
Thornton, M., Patterson, C., Lal, P., Burford, N., Yue, H., Gandhi, A.R.,
Elliot, V.S., Ramkumar, J., Baughn, M.R., Kallick, D.A., Wallia, N.K.,
Hafalia, A.U., Yao, M.G., Lu, Y., Tribouley, C.M., Policky, J.L.,
Kearney, L., Graul, R.C., Warren, B.A. and Ding, L.
G-Protein coupled receptors
Patent: WO 0210387-A 34 07-FEB-2002;
Incyte Genomics, Inc. (US)
FEATURES
Location/Qualifiers
1..1458
/mol_type="unassigned DNA"
source
/organism="Homo sapiens"
```

```
QY 5 ATGAGGCGCGTGTAGCTTTTCAAGTGGCCACCGCGCGCTTGGCTTGGCTTGGAGCCGAG 64
Db 1 ATGAGGCGCGTGTAGCTTTTCAAGTGGCCACCGCGCGCTTGGCTTGGCTTGGAGCCGAG 60
QY 65 ACCAGACGAGACCGCGGACCCAGCCAGCCAGAGAGGATCTCGTTCGACCCCGAGCGGC 124
Db 61 ACCAGCAG-----CGGACCCAGCCAGAGAGGATCTCGTTCGACCCCGAGCGGC 114
QY 125 GCGGTCTGCGGCGCGAGGGCGCGCTTCTGTTTCAAGTGGCTTGGCTTGGCTTGGAGCTG 184
Db 115 GCGGTCTGCGGCGCGAGGGCGCGCTTCTGTTTCAAGTGGCTTGGCTTGGCTTGGAGCTG 174
QY 185 CTAGTGTCTGATCGCTGCCACTTTCTGTGGAACTGCTGTTCCGGTCAACCATCCCG 244
Db 175 CTAGTGTCTGATCGCTGCCACTTTCTGTGGAACTGCTGTTCCGGTCAACCATCCCG 234
QY 245 CGGCTCGTGTCTTCCACCGCGTGGCGATAACTTTGGTGGCTCGACGGCGCTCTCGGAC 304
Db 235 CGGCTCGTGTCTTCCACCGCGTGGCGATAACTTTGGTGGCTCGACGGCGCTCTCGGAC 294
QY 305 GAACTAGTGGACGCTGCGGATGCCAGCGCTGGAGTGGAGTGGAGTGGAGTGGAGTGGAG 364
Db 295 GAACTAGTGGACGCTGCGGATGCCAGCGCTGGAGTGGAGTGGAGTGGAGTGGAGTGGAG 354
QY 365 CGTGGCTGCTGGCGCGGAGCTGTGCCACAGTGTGGATCTCTCTCGACCGCGAGCTGT 424
Db 355 CGTGGCTGCTGGCGCGGAGCTGTGCCACAGTGTGGATCTCTCTCGACCGCGAGCTGT 405
QY 425 CTGTGCTGCTGGCGCGGCTCGGAAAGTGGCGGCTATCGCTTGGCGCGAGCGCGGCGC 484
Db 406 CTGTGCTGCTGGCGCGGCTCGGAAAGTGGCGGCTATCGCTTGGCGCGAGCGCGGCGC 465
QY 485 ATCACACGACCTGACGACACAGCTGGCACCGCGCGCGCGCTGCTGCTGCTGCTGCTG 544
Db 466 ATCACACGACCTGACGACACAGCTGGCACCGCGCGCGCGCTGCTGCTGCTGCTGCTG 525
QY 545 GCGCTCGCGCGGCTCGCGCTCATCGCGCTCGCGCTCATCGCGCTCGCGCTCGCGCTCG 604
Db 526 GCGCTCGCGCGGCTCGCGCTCATCGCGCTCATCGCGCTCATCGCGCTCGCGCTCGCGCT 585
QY 605 GAGGTGTGCGACGCTCGGCTCAGCGCTGCGAGTGGAGCGCGGAGACCTCTATGCGCGC 664
Db 586 GAGGTGTGCGACGCTCGGCTCAGCGCTGCGAGTGGAGCGCGGAGACCTCTATGCGCGC 645
QY 665 TTCTCCACCGCGCGCTTCCACCTGCGCTTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 724
Db 646 TTCTCCACCGCGCGCTTCCACCTGCGCTTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 705
QY 725 ATCTAGAGCGCGCAAGTTTCTGTTTGGCGCGCGCGCGAGAGCTGCTGCTGCTGCTGCTG 784
Db 706 ATCTAGAGCGCGCAAGTTTCTGTTTGGCGCGCGCGAGAGCTGCTGCTGCTGCTGCTGCT 765
QY 785 GGCACCATGAGGTGAGTCCAGAGTAAAGAGACCTGATGAGGTGAGGTGAGGTGAGTTC 844
Db 766 GGCACCATGAGGTGAGTCCAGAGTAAAGAGACCTGATGAGGTGAGGTGAGTTC 813
QY 845 ACGGACATTTGCAAGCAACGCTGCTTCCAGGTGAGCGGAGCTCTCTGGCGGAGCAG 904
Db 814 ACGGACATTTGCAAGCAACGCTGCTTCCAGGTGAGCGGAGCTCTCTGGCGGAGCAG 873
QY 905 AAGGAGAGCGGAGCGGACCATGATGTTGGAAATTTGATTTGGCTGTTTGTGCTGCTG 964
Db 874 AAGGAGAGCGGAGCGGACCATGATGTTGGAAATTTGATTTGGCTGTTTGTGCTGCTG 933
QY 965 ATCCCTTCTTCTGCTGAGGAACTCATGACCCACTCTGCTGCTGCGAGCTGCCCCCATC 1024
```

Db	934	ATCCCCCTTCTGACGGAATCATCAGCCCATCTGTGCTCGACGCTGCCCCCATC	993	Db	355	CGTCGCTGCTGGCGCGGAGCCTGTGCCACGTGTGATCTCTCTTCACG	-----TG 405
Qy	1025	TGGAAGAAGCATATTTCTGTGGCTTGCTACTCAATTTCTTCTCAACCCCTGATTAC	1084	Qy	425	CTGTGCTGCCCGCGCGGCTTCGGGAACGTGGCGGCAATCGCCCTGGCGCGGAGCC	484
Db	994	TGGAAGAAGCATATTTCTGTGGCTTGCTACTCAATTTCTTCAACCCCTGATTAC	1053	Db	406	CTGTGCTGCCCGCGCGGCTTCGGGAACGTGGCGGCAATCGCCCTGGCGCGGAGCC	465
Qy	1085	ACAGCTTTTAAACAAGAACTACAAATGCTTTCAAGAGCCTCTTTACTAAGCAGATGA	1144	Qy	485	ATCACACGACCTCGACGACACGCTGGCGACCCCGCAGCGCGGCTCTGTCTCATGATC	544
Db	1054	ACAGCTTTTAAACAAGAACTACAAATGCTTTCAAGAGCCTCTTTACTAAGCAGATGA	1113	Db	466	ATCACACGACCTCGACGACACGCTGGCGACCCCGCAGCGCGGCTCTGTCTCATGATC	525
Qy	1145	ACACAGGG 1152		Qy	545	GGCTCGCCCGGGTGCCTCGGCGCTCATCGCCCTCGCGCCGCTCTTTTGGCGGGGC	604
Db	1114	ACACAGGG 1121		Db	526	GGCTCACCCGGTGCCTCGGCGCTCATCGCCCTCGCGCCGCTCTTTTGGCGGGGC	585
RESULT 5				Qy	605	GAGGTGTGGAGCTCGGCTCCAGCGCTCCAGGTGAGCGGGAACCTCTATGCGCC	664
AX342670				Db	586	GAGGTGTGGAGCTCGGCTCCAGCGCTCCAGGTGAGCGGGAACCTCTATGCGCC	645
LOCUS	AX342670	1119 bp	DNA	linear	PAT 12-JAN-2002		
DEFINITION	Sequence 25 from Patent WO0198351.			Qy	665	TTCTCACCCCGCGGCTTCACCTGCGCTTCACCTGCGCTGGCGCTCTTTTGGCGGGAG	724
ACCESSION	AX342670			Db	646	TTCTCACCCCGCGGCTTCACCTGCGCTTCACCTGCGCTGGCGCTCTTTTGGCGGGAG	705
VERSION	AX342670.1	GI:18152050		Qy	725	ATCTACGAGCGGCAAGTTTCTGTTTCGCGCGCCCGCGAGAGCTGTGCTGCGCTTGGCG	784
KEYWORDS	Homo sapiens (human)			Db	706	ATCTACGAGCGGCAAGTTTCTGTTTCGCGCGCCCGCGAGAGCTGTGCTGCGCTTGGCG	765
SOURCE	Homo sapiens			Qy	785	GCCACCATGAGGTGAGTCCAAAGTAAAGAACACCTGATGAGGCTGAAGTGTGTTTC	844
ORGANISM	Homo sapiens			Db	766	GCCACCATGAGGTGAGTCCAAAGTAAAGAACACCTGATGAGGCTGAAGTGTGTTTC	819
REFERENCE				Qy	845	ACGGCACATTCGAAAGAACCGTGTCTTCCAGGTGAGCGGGGACTCTTGGCGGGAGAG	904
AUTHORS	Lal, P., Baughn, M.R., Hafalia, A.J., Nguyen, D.B., Gandhi, A.R.,			Db	820	ACGGCACATTCGAAAGAACCGTGTCTTCCAGGTGAGCGGGGACTCTTGGCGGGAGAG	879
	Kallick, D.A., Griffin, J.A., Yue, H., Khan, F.A., Patterson, C.,			Qy	905	AAGGAGAGGCGAGCAGCAGCATGATGTTGGGAATTCGATTGGCGCTGTTGCTGTGCTGG	964
	Lu, D.A., Tribolet, C.M., Lu, Y., Wallia, N.K., Graul, R., Yao, M.G.,			Db	880	AAGGAGAGGCGAGCAGCAGCATGATGTTGGGAATTCGATTGGCGCTGTTGCTGTGCTGG	939
	Yang, J., Ramkumar, J., Au-Young, J., Hernandez, R., Walsh, R.T. and			Qy	965	ATCCCTTCTTCTGAGGAACTCATCAGCCCACTCTGTGCTGAGCTGCCCCCATC	1024
	Borowsky, M.L.			Db	940	ATCCCTTCTTCTGAGGAACTCATCAGCCCACTCTGTGCTGAGCTGCCCCCATC	999
JOURNAL	Patent: WO 0198351-A 25 27-DEC-2001;			Qy	1025	TGGAAGAAGCATATTTCTGTGGCTTGCTACTCAATTTCTTCTCAACCCCTGATTAC	1084
	Incyte Genomics, Inc. (US)			Db	1000	TGGAAGAAGCATATTTCTGTGGCTTGCTACTCAATTTCTTCTCAACCCCTGATTAC	1059
FEATURES				Qy	1085	ACAGCTTTTAAACAAGAACTACAAATGCTTTCAAGAGCCTCTTTACTAAGCAGATGA	1144
source	1..1119			Db	1060	ACAGCTTTTAAACAAGAACTACAAATGCTTTCAAGAGCCTCTTTACTAAGCAGATGA	1119
	/organism="Homo sapiens"			RESULT 6			
	/mol_type="unassigned DNA"			AX526745			
	/db_xref="taxon:9606"			LOCUS	AX526745		
	/note="Incyte ID No: 7474977CB1"			DEFINITION	Sequence 29 from Patent WO0224733.		
ORIGIN				ACCESSION	AX526745		
Query Match	92.0%;	Score 1060;	DB 6;	VERSION	AX526745.1	GI:25171536	
Best Local Similarity	97.7%;	Pred. No. 5.9e-167;	Length 1119;	KEYWORDS	unidentified		
Matches 1114;	Conservative 0;	Mismatches 5;	Indels 21;	SOURCE	unidentified		
			Gaps 3;	ORGANISM	unclassified.		
Qy	5	ATGAGGCGCGTAGCCCTTCAGTGGCCACCGCGGCTTGCCCTTGCCCTGGGACCCGAG	64	REFERENCE	1		
Db	1	ATGAGGCGCGTAGCCCTTCAGTGGCCACCGCGGCTTGCCCTTGCCCTGGGACCCGAG	60	AUTHORS	Mishra, V.S., Spytek, K.A., Taupier, R.J., Vernet, C.A., Colman, S.D.,		
Qy	65	ACCAGCAGCGGACCCCGGACCCCAAGCCCGAGAGGGATCTCGGTTCCGACCCCGAGCGGC	124		Gorman, L., Tchernev, V.T., Malyankar, U.M., Shenoy, S., Tchernev, V.T.,		
Db	61	ACCAGCAG-----CGGAGACCCCAAGCCCGAGAGGGATCTCGGTTCCGACCCCGAGCGGC	114		Padigar, M., Patturajan, M., Burgess, C.E., Smithson, G., Millet, I.,		
Qy	125	GCGGCTCTGCGGCGGAGCGGCGCCCTTCTGTTCTTCAAGGTCCTGTTGGTGGAGCCTG	184		Peyman, J.A., Stone, D., Gunther, E. and Ellerman, K.		
Db	115	GCGGCTCTGCGGCGGAGCGGCGCCCTTCTGTTCTTCAAGGTCCTGTTGGTGGAGCCTG	174		Human polynucleotides and polypeptides encoded thereby		
Qy	185	CTAGTGTGCTGATCGTGCACCTTCTGTTGAACTGTTGTTGTTGTTGTTGTTGTTGTTG	244		Patent: WO 0224733-A 29 28-MAR-2002;		
Db	175	CTAGTGTGCTGATCGTGCACCTTCTGTTGAACTGTTGTTGTTGTTGTTGTTGTTGTTG	234		Curagen Corporation (US)		
Qy	245	CGGTCCTGCTTCCACCGCGTGCAGCAATCTTGGTGGCCTCGAGCGGCTCGGAC	304		Location/Qualifiers		
Db	235	CGGTCCTGCTTCCACCGCGTGCAGCAATCTTGGTGGCCTCGAGCGGCTCGGAC	294		1..1155		
Qy	305	GAATAGTGACGCTGCGGATGCCACCGAGCTGCGGAGTGTGTTGCGACCGGCGGA	364		/organism="unidentified"		
Db	295	GAATAGTGACGCTGCGGATGCCACCGAGCTGCGGAGTGTGTTGCGACCGGCGGA	354		/mol_type="unassigned DNA"		
Qy	365	CGTGGCTGTGGCGGAGCCTGTGCCAGTGTGATCTCTTTCAGCGCGGCGGCTGT	424				

```
/db xref="taxon:32644"
/note="Description of Unknown Organism: NOVX Nucleic Acid"
```

ORIGIN

Query Match	88.2%	Score 1016	DB 6	Length 1155
Best Local Similarity	97.2%	Pred. No. 1.3e-159		
Matches 1087	Conservative	0	Mismatches	Indels 31
QY	36	CCGGCGTTGCCCTTGGCCCTGGGACCCAGACACAGCAGC- GGACCCCGGGACCCCAAGCCCG 94		
Db	68	CCGGCGTTGCCCTTGGCCCTGGGACCCAGACACAGCAGCAGGACCCGCGGACCCCAAGCCCG 127		
QY	95	AGAGGATATCTCGTTTCGACCCCGAGCGCGCGCTCTCTGCGGGCCGAGGGCGGCCCTTC 154		
Db	128	AGAGGATATCTCGTTTCGACCCCGAGCGCGCGCTCTCTGCGGGCCGAGGGCGGCCCTTC 187		
QY	155	TCCTGCTTTCACGCTCTGCTGGTGGTGGCTGCTAGTCTGCTGATCGCTCCCTCCCTTCCTG 214		
Db	188	TCCTGCTTTCACGCTCTGCTGGTGGTGGCTGCTAGTCTGCTGATCGCTCCCTCCCTTCCTG 247		
QY	215	TGGAACCTGCTGGTTCCGGTTCACCATCCCGCGGTCCGTGCCCTTCACCGCGTCCGAT 274		
Db	248	TGGAACCTGCTGGTTCCGGTTCACCATCCCGCGGTCCGTGCCCTTCACCGCGTCCGAT 307		
QY	275	AACCTTGGTGCCTCGACCGCGCTCTCGGACGAATAGTGGCAGCGCTGGCGATGCCACCG 334		
Db	308	AACCTTGGTGCCTCGACCGCGCTCTCGGACGAATAGTGGCAGCGCTGGCGATGCCACCG 367		
QY	335	AGCCTGCGAGTAGCTGTTCGACCGGCGCAGCTCGGCTGTGGCGCGGAGCCCTGTGCCAC 394		
Db	368	AGCCTGCGAGTAGCTGTTCGACCGGCGCAGCTCGGCTGTGGCGCGGAGCCCTGTGCCAC 418		
QY	395	GTGTGGATCTCTTCGACCGCGGAGCCTGTGTGTGCTGCCCGCGGCTCGGGACGCTG 454		
Db	419	GTGTGGATCTCTCTTCGACGCGC-----CTGTGCTGCCCGCGGCTCGGGACGCTG 469		
QY	455	GCGGCCATCGCCTTGGCGCGGACGCGGCGCATCACAGCGACCTTCGACGACACGCTGCGC 514		
Db	470	GCGGCCATCGCCTTGGCGCGGACGCGGCGCATCACAGCGACCTTCGACGACACGCTGCGC 529		
QY	515	ACCGCAGCGCGCTCTGTGCTCATGATCGCGCTCCCGCGGCTCGGCGCTCATC 574		
Db	530	ACCGCAGCGCGCTCTGTGCTCATGATCGCGCTCCCGCGGCTCGGCGCTCATC 589		
QY	575	GCCCTGCGCGCTGCTCTTTGGCGCGGCGAGGTGTGGACGCTCGGCTCCAGCGCTGC 634		
Db	590	GCCCTGCGCGCTGCTCTTTGGCGCGGCGAGGTGTGGACGCTCGGCTCCAGCGCTGC 649		
QY	635	CAGGTGACCGGGAAACCTCTATGCGCGCTTCCTCCACCGCGCGGCTTCACCTCCG 694		
Db	650	CAGGTGACCGGGAAACCTCTATGCGCGCTTCCTCCACCGCGCGGCTTCACCTCCG 709		
QY	695	CTTGGCGTGGTCCGTTTGTCTACCGGAAGATCTACGAGGCGGCGAATTCGTTTCGGT 754		
Db	710	CTTGGCGTGGTCCGTTTGTCTACCGGAAGATCTACGAGGCGGCGAATTCGTTTCGGT 769		
QY	755	GCGCCCGGAGCTGTGCTCGCGTTCGCGCCACCATGACAGGTGAGGTCCCAAGGTAAG 814		
Db	770	GCGCCCGGAGCTGTGCTCGCGTTCGCGCCACCATGACAGGTGAGGTCCCAAGGTAAG 817		
QY	815	GAAGCACCTGATGAGGCTGAAGTGTGTTCACGCGCATATTGCAAGCAACGGTTCCTTC 874		
Db	818	GAAGCACCTGATGAGGCTGAAGTGTGTTCACGCGCATATTGCAAGCAACGGTTCCTTC 877		
QY	875	CAGGTGACGGGACTCTCTGCGGAGCAGAGGAGGCGGAGCGACCATGATGGTGGGA 934		
Db	878	CAGGTGACGGGACTCTCTGCGGAGCAGAGGAGGCGGAGCGACCATGATGGTGGGA 937		
QY	935	ATTCTGATTGGCGTGTGTGTGCTGTGGATCCCGCTTCTTCTTCTGACGGAATCATCAGC 994		
Db	938	ATTCTGATTGGCGTGTGTGTGCTGTGGATCCCGCTTCTTCTTCTGACGGAATCATCAGC 997		
QY	995	CCACTCTGTGCTGAGCGCTGCCCGCCCATCTGGAAAAAGATATTTCTGTGGTTGGCTAC 1054		

Db	998	CCACTCTGTGCGTGCAGCCTGCCCCCACTCTGGAAGAAGCATATTTCTGTGGCTTGGCTAC	1057
Qy	1055	TCCAAATCTTTCTTCAACCCCTGATTACACAGCTTTTAAACAAGAAGTACACAATGCC	1114
Db	1058	TCCAAATCTTTCTTCAACCCCTGATTACACAGCTTTTAAACAAGAAGTACACAATGCC	1117
Qy	1115	TTCAAGAGCCTCTTTTACTTAAGCAGAGATGAACACAGGG	1152
Db	1118	TTCAAGAGCCTCTTTTACTTAAGCAGAGATGAACACAGGG	1155
RESULT 7			
AX646417			
LOCUS	AX646417	3086 bp	DNA linear PAT 04-MAR-2003
DEFINITION	Sequence 609 from Patent EP1270724.		
ACCESSION	AX646417		
VERSION	AX646417.1	GI:28798798	
KEYWORDS			
SOURCE	Homo sapiens (human)		
ORGANISM	Homo sapiens		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.		
AUTHORS	1 Suwa, M., Asai, K., Akiyama, Y. and Aburatani, H.		
TITLE	Guanosine triphosphate-binding protein coupled receptors		
JOURNAL	Patent: EP 1270724-A 609 02-JAN-2003;		
	National Institute of Advanced Industrial Science and Technology		
	(JPN); Center for Advanced Science and Technology Incubation, Ltd.		
	(JPN)		

FEATURES

```

source
Location/Qualifiers
1. .3086
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
3086(201 277 1118. 2192.2218. .2886)

```

ORIGIN

	Query Match	65.6%	Score 756.2	DB 6	Length 3086
	Best Local Similarity	95.4%	Prod. No. 2.1e-116		
	Matches 810	Conservative	0	Mismatches 13	Indels 26
	Gaps				
Qy	1	CGCCATGGAGGCGCGCTAGCCCTTTTCAGTGGCCACCGCGCGCTTGCCTTGCCTTGCCTGGGACC	60		
Db	1798	CGCCATGGAGGCGCGCTAGCCCTTTTCAGTGGCCACCGCGCGCTTGCCTTGCCTTGCCTGGGACC	185		
Qy	61	CGAGACGACGAGC-GGAGCCCGGGACCCCAAGCCCGAGAGGAGTACTCGGTTGCAGCCCGA	119		
Db	1858	CGAGACGACGACGAGGACCCCGGGACCCCAAGCCCGAGAGGAGTACTCGGTTGCAGCCCGA	191		
Qy	120	CGCGCGCGCTCTGCGCGGGCGGAGGGCGCGCCCTTCTCTGTCTTCAACGGTCTCTGTGGTGA	179		
Db	1918	CGCGCGCGCTCTGCGGGCGGAGGGCGCGCCCTTCTCTGTCTTCAACGGTCTCTGTGGTGA	197		
Qy	180	CGCTGCTAGTGTGCTGTATGCGTGGCACTTTCCTGTGAAACTGTGGTTCGGGTCA	239		
Db	1978	CGCTGCTAGTGTGCTGTATGCGTGGCACTTTCCTGTGAAACTGTGGTTCGGGTCA	203		

240 TCCCGCGGTCGGTCCCTTCCACCGCGTCCGCGATAACTTGGTGGCTCGAGCGCGTCT 299
 Db TCCCGCGGTCGGTCCCTTCCACCGCGTCCGCGATAACTTGGTGGCTCGAGCGCGTCT 2097
 QY 300 CGGACGAATAGTGGCAGCGTGGCGATGCCACCGAGCTGGCGAGTGGAGTGTGACCG 359
 Db CGGACGAATAGTGGCAGCGTGGCGATGCCACCGAGCTGGCGAGTGGAGTGTGACCG 2157
 QY 360 GCGGACGTCGCTGCTGGCGCGGAGCCGTGTGCCAGTGTGATCTCCCTGACGCGCGAG 419
 Db GCGGACGTCGCTGCTGGCGCGGAGCCGTGTGCCAGTGTGATCTCCCTGACGCGCGAG 2217
 QY 420 CTTCT-----CTGTGCTGCCCGCGCGCGCTCGGGAACGTG 454
 Db CTTGTGCCACGTGTGATCTCTTCCACGGCTGTGTCTCCCGCGCGCTCGGGAACGTG 2277
 QY 455 GCGGCATCGCTGGCGCGGAGCGGCGATCAACGGGACCTGACGACACAGCTGGCG 514
 Db GCGGCATCGCTGGCGCGGAGCGGCGATCAACGGGACCTGACGACACAGCTGGCG 2337
 QY 515 ACCGCGAGCGCGCTGCTTGTGGCGCGGCGAGTGTGCGACGCTCGGCTCCAGCGTGC 574
 Db ACCGCGAGCGCGCTGCTTGTGGCGCGGCGAGTGTGCGACGCTCGGCTCCAGCGTGC 2397
 QY 575 GCCCTCGCGCGCTGCTTGTGGCGCGGCGAGTGTGCGACGCTCGGCTCCAGCGTGC 634
 Db GCCCTCGCGCGCTGCTTGTGGCGCGGCGAGTGTGCGACGCTCGGCTCCAGCGTGC 2457
 QY 635 CAGGTGAGCGGGAACCTCTATGCGCGCTTCCACCGCGCGCTTCCACCTGCGG 694
 Db CAGGTGAGCGGGAACCTCTATGCGCGCTTCCACCGCGCGCTTCCACCTGCGG 2517
 QY 695 CTTGGCGTGGTGGCTTGTCTACGGAGATCTACGAGGCGGCGCAAGTTCGTTTGGC 754
 Db CTTGGCGTGGTGGCTTGTCTACGGAGATCTACGAGGCGGCGCAAGTTCGTTTGGC 2577
 QY 755 CGCGCGGAGAGTGTGCTGCTGCGCTTGGCGCGACCATGAGTGTGAGTCCAGGTAAG 814
 Db CGCGCGGAGAGTGTGCTGCTGCGCTTGGCGCGACCATGAGTGTGAGTGTGAGG 2637
 QY 815 GAGCACT 823
 Db AACGTTGCT 2646

RESULT 8
 AB065453
 LOCUS Homo sapiens gene for seven transmembrane helix receptor, linear PRI 23-JUL-2002
 DEFINITION isolate:CBRC7TM_16.
 ACCESSION AB065453
 VERSION AB065453.1 GI:21928219
 KEYWORDS
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1
 AUTHORS Suwa,M., Sato,T., Okouchi,I., Arita,M., Futami,K., Matsumoto,S., Tezumi,S., Aburatani,H., Asai,K. and Akiyama,Y.
 TITLE Genome-wide discovery and analysis of human seven transmembrane helix receptor genes
 JOURNAL Unpublished
 REFERENCE 2 (bases 1 to 44312)
 AUTHORS Suwa,M.
 TITLE Direct Submission
 JOURNAL Submitted (11-JUL-2001) Makiko Suwa, Computational Biology Research Center (CBRC), National Institute of Advanced Industrial Science and Technology (AIST); 2-41-6 Aomi Koto-ku, Tokyo 135-0064, Japan (E-mail:m-suwa@aist.go.jp, URL:http://www.cbrc.jp/, Tel:81-3-3599-8080, Fax:81-3-3599-8081)
 COMMENT This sequence is a seven transmembrane helix receptor candidate

predicted from the whole human genome sequences using our automated system that contains programs of gene finding (GenePredictor), sequence search, motif-domain assignment and transmembrane helix prediction.

And the sequence is submitted by the collaborative project between [Computational Biology Research Center (CBRC), National Institute of Advanced Industrial Science and Technology (AIST)] and [Genome Science Division, Research Center for Advanced Science and Technology (RCAST), University of Tokyo].

FEATURES

source
 1. 44312
 /organism="Homo sapiens"
 /mol_type="genomic DNA"
 /isolate="CBRC7TM_16"
 /db_xref="taxon:9606"
 /chromosome="2"
 misc_feature
 join(201..1018,43777..44112)
 /product="seven transmembrane helix receptor"
 /evidence=not_experimental

ORIGIN

Query Match 65.6%; Score 756.2; DB 9; Length 44312;
 Best Local Similarity 95.4%; Pred. No. 1.2e-116;
 Matches 810; Conservative 0; Mismatches 13; Indels 26; Gaps 2;

QY 1 CGCCATGAGGCGCGTAGCCCTTTCAGTGGCCACCGCGCGTTCGCCCTGGCGGACC 60
 Db 197 CGCCATGAGGCGCGCTAGCCCTTTCAGTGGCCACCGCGCGTTCGCCCTGGCGGACC 256

QY 61 CGAGACAGCAGC-GGACCCCGGACCCCAAGCCCGAGAGGATACTCGGTTTCGACCCCGA 119
 Db 257 CGAGACAGCAGCAGGACCCCGGACCCCAAGCCCGAGAGGATACTCGGTTTCGACCCCGA 316

QY 120 GCGGCGCGTCTCGCGGCGCGAGGCGCGCTTCTCTGCTTTCACGCTCTCGTGTGA 179
 Db 317 GCGGCGCGTCTCGCGGCGCGAGGCGCGCTTCTCTGCTTTCACGCTCTCGTGTGA 376

QY 180 CGCTCTAGTGTCTGCTGATGCTGCGACTTCTCTGTTGAGAACTGCTGGTTCGGTACCA 239
 Db 377 CGCTCTAGTGTCTGCTGATGCTGCGACTTCTCTGTTGAGAACTGCTGGTTCGGTACCA 436

QY 240 TCCCGCGGTCGGTCCCTTCCACCGCGTCCGCGATAACTTGGTGGCTCGAGCGCGTCT 299
 Db 437 TCCCGCGGTCGGTCCCTTCCACCGCGTCCGCGATAACTTGGTGGCTCGAGCGCGTCT 496

QY 300 CGGACGAATAGTGGCAGCGTGGCGATGCCACCGAGCTGGCGAGTGGAGTGTGACCG 359
 Db 497 CGGACGAATAGTGGCAGCGTGGCGATGCCACCGAGCTGGCGAGTGGAGTGTGACCG 556

QY 360 GCGGACGTCGCTGCTGGCGCGGAGCTGTGCCAGTGTGATCTCTTTCGACGCGGAG 419
 Db 557 GCGGACGTCGCTGCTGGCGCGGAGCTGTGCCAGTGTGATCTCTTTCGACGCGGAG 616

QY 420 CTTCT-----CTGTGCTGCCCGCGCGCGCTCGGGAACGTG 454
 Db 617 CTTGTGCCACGTGTGATCTCTTTCACGGCTGTGTCTCCCGCGCGCTCGGGAACGTG 676

QY 455 GCGGCATCGCTGGCGCGGAGCGGCGATCAACGGGACCTGACGACACAGCTGGCG 514
 Db 677 GCGGCATCGCTGGCGCGGAGCGGCGATCAACGGGACCTGACGACACAGCTGGCG 736

QY 515 ACCGCGAGCGCGCTGCTTGTGGCGCGGAGTGTGCGACGCTCGGCTCGGCGCTATC 574
 Db 737 ACCGCGAGCGCGCTGCTTGTGGCGCGGAGTGTGCGACGCTCGGCTCGGCGCTATC 796

QY 575 GCCTCGCGCGCTGCTTGTGGCGCGGAGTGTGCGACGCTCGGCTCCAGCGTGC 634
 Db 797 GCCTCGCGCGCTGCTTGTGGCGCGGAGTGTGCGACGCTCGGCTCCAGCGTGC 856

QY 635 CAGGTGAGCGGGAACCTCTTATGCGCGCTTCTCCACCGCGCGCGCTTCCACCTGCGC 694
 Db 857 CAGGTGAGCGGGAACCTCTTATGCGCGCTTCTCCACCGCGCGCGCTTCCACCTGCGC 916

```

695 CTTGCGGTGGTGGCGTTGTCTACGGAGATCTACGAGCGCGCAAGTTCTGTTTCGGC 754
|||||
917 CTTGCGGTGGTGGCGTTGTCTACGGAGATCTACGAGCGCGCAAGTTCTGTTTCGGC 976
|||||
755 CGCGCGCGAGAGCTGCTCGGTGCGCGCCACCATGCGAGGTGAGTCCAAAGTAAAG 814
|||||
977 CGCGCGCGAGAGCTGCTCGGTGCGCGCCACCATGCGAGGTGAGGTGGGCTGAGG 1036
|||||

QY 815 GAAGCACT 823
DB 1037 AACGTTGCT 1045

RESULT 9
LOCUS AC009404 112883 bp DNA linear PRI 07-OCT-2000
DEFINITION Homo sapiens BAC clone RP11-28H22 from 2, complete sequence.
ACCESSION AC009404
VERSION AC009404.5 GI:10716633
KEYWORDS HTG
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
1 (bases 1 to 112883)
Sulston, J.E. and Waterston, R.
Toward a complete human genome sequence
Genome Res. 8 (11), 1097-1108 (1998)
99063792
9847074
REFERENCE 2 (bases 1 to 112883)
AUTHORS Cordes, M., Walker, C., Stoneking, T. and Reitz, L.
TITLE The sequence of Homo sapiens BAC clone RP11-28H22
JOURNAL Unpublished
REFERENCE 3 (bases 1 to 112883)
AUTHORS Waterston, R.H.
TITLE Direct Submission
JOURNAL Submitted (21-AUG-1999) Genome Sequencing Center, Washington
University School of Medicine, 4444 Forest Park Parkway, St. Louis,
MO 63108, USA
4 (bases 1 to 112883)
AUTHORS Waterston, R.H.
TITLE Direct Submission
JOURNAL Submitted (26-JUL-2000) Genome Sequencing Center, Washington
University School of Medicine, 4444 Forest Park Parkway, St. Louis,
MO 63108, USA
5 (bases 1 to 112883)
AUTHORS Waterston, R.
TITLE Direct Submission
JOURNAL Submitted (07-OCT-2000) Department of Genetics, Washington
University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA
On Oct 7, 2000 this sequence version replaced gi:9454621.
----- Genome Center
Center: Washington University Genome Sequencing Center
Web site: http://genome.wustl.edu/gsc
Contact: sapiense@wustl.wustl.edu
----- Summary Statistics
-----
Center project name: H_NH0028H22
-----

```

NOTICE: This sequence may not represent the entire insert of this clone. It may be shorter because we only sequence overlapping clone sections once, or longer because we provide a small overlap between neighboring data submissions.

This sequence was finished as follows unless otherwise noted: all regions were double stranded, sequenced with an alternate chemistry, or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by sequence from more than one subclone; and the assembly was confirmed by restriction digest.

MAPPING INFORMATION:

Mapping information for this clone was provided by Dr. John D. McPherson, Department of Genetics, Washington University, St. Louis MO. For additional information about the map position of this sequence, see <http://genome.wustl.edu/gsc>

SOURCE INFORMATION:

The RPC1-11 human BAC library was made from the blood of one male donor, as described by Osoegawa, K., Woon, P. Y., Zhao, B., Frengen, E., Taten, M., Catanese, J. J. and de Jong, P. J. (1998) An improved approach for construction of bacterial artificial chromosome libraries. Genomics 51:1-8. The clone may be obtained either from Research Genetics, Inc. (<http://www.resgen.com>) or Pieter de Jong and coworkers at the Roswell Park Cancer Institute (<http://bacpac.med.buffalo.edu>)

NEIGHBORING SEQUENCE INFORMATION:

The clone sequenced to the left is RP11-425F6, 200 bp overlap; the clone sequenced to the right is RP11-98C1, 200 bp overlap. Actual start of this clone is at base position 157644 of RP11-425F6; actual end is at base position 21685 of RP11-98C1.

The sequence RP11-28H22 contains imperfect dinucleotide (CT and GC) repeats from base position 53407 to 53507. The region is covered by a single clone; the fidelity of the sequence cannot be guaranteed. The length of sequence is consistent with PCR from clone DNA, and corresponds to restriction digest information: hindIII band size 5664 in silico, and 5694 real.

There are polymorphic base pair differences in the overlaps between the clones RP11-425F6, RP11-28H22, and RP11-425F6.

FEATURES	Source
1..112883	/organism="Homo sapiens"
	/mol_type="Genomic DNA"
	/db_xref="taxon:9606"
	/chromosome="2"
	/map="2"
	/clone="RP11-28H22"
	/clone_lib="RPC1-11"
misc_feature	1..101
misc_feature	/note="similar to EST AW069620 (NID:g6024618) cr47g08.xl"
misc_feature	1..101
misc_feature	/note="similar to EST AW803315 (NID:g7855185)"
misc_feature	8..101
misc_feature	/note="similar to EST AA108891 (NID:g1660325) ml56c09.r1"
misc_feature	56..101
misc_feature	/note="similar to EST AA870919 (NID:g2966364) vq28a09.r1"
repeat_region	644..733
repeat_region	/rpt_family="MER1_type"
repeat_region	734..807
repeat_region	/rpt_family="MER1_type"
repeat_region	1062..1194
repeat_region	/rpt_family="Alu"
misc_feature	2022..2165
misc_feature	/note="similar to EST AW069620 (NID:g6024618) cr47g08.xl"
misc_feature	2022..2165
misc_feature	/note="similar to EST AW803315 (NID:g7855185)"
misc_feature	2034..2165
misc_feature	/note="similar to EST AA108891 (NID:g1660325) ml56c09.r1"
misc_feature	2034..2165
misc_feature	/note="similar to EST AA870919 (NID:g2966364) vq28a09.r1"
misc_feature	2043..2165
misc_feature	/note="similar to EST AW582239 (NID:g7257288)"
misc_feature	2044..2165
misc_feature	/note="similar to EST AA870844 (NID:g2966289) vq27b06.r1"
misc_feature	2044..2165
misc_feature	/note="similar to EST A1021272 (NID:g3235608) ub04e12.r1"
misc_feature	2084..2165
misc_feature	/note="similar to EST AW892944 (NID:g8057149)"
repeat_region	2706..2819
repeat_region	/rpt_family="MER81"

[illegible]

REFERENCE
AUTHORS
TITLE
JOURNAL
REFERENCE
AUTHORS

Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
1 (bases 1 to 184328)
Birren, B., Linton, L., Nusbaum, C., Lander, E., Allen, N., Anderson, M.,
Baldwin, J., Barna, N., Beckerly, R., Boguslavsky, L., Boukhgalter, B.,
Brown, A., Casle, A., Colangelo, M., Collins, S., Collymore, A.,
Cooke, P., DeArrellano, K., Dewar, K., Domino, M., Donelan, L., Doyle, M.,
Ferreira, P., FitzHugh, M., Forrest, C., Funke, R., Gage, D.,
Galagan, J., Gardyna, S., Grant, G., Hagos, B., Heaford, A., Horton, L.,
Holland, J. C., Johnson, R., Jones, C., Kann, L., Karatas, A., Klein, J.,
Leboczky, J., Lieu, C., Locke, K., Macdonald, P., Marquis, N.,
McEwan, P., McMurk, A., McKernan, K., McLaughlin, J., O'Donnell, J.,
Morrow, J., Naylor, J., Norman, C. H., O'Connor, T., O'Donnell, P.,
Peterson, K., Pollara, V., Riley, R., Roy, A., Santos, R., Severy, P.,
Stange-Thomann, N., Stojanovic, N., Subramanian, A., Talamas, J.,
Tesfaye, S., Tirrell, A., Vassiliev, H., Vo, A., Wheeler, J., Wu, X.,
Wymann, D., Ye, W. J., Zimmer, A. and Zody, M.

Direct Submission
Submitted (08-OCR-1999) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
On Mar 1, 2000 this sequence version replaced gi:6957833.
All repeats were identified using RepeatMasker:
Smit, A. F. A. & Green, P. (1996-1997)
<http://ftp.genome.washington.edu/RM/RepeatMasker.html>

----- Genome Center
Center: Whitehead Institute/ MIT Center for Genome Research
Center code: WIBR
Web site: <http://www-seq.wi.mit.edu>
Contact: sequence_submissions@genome.wi.mit.edu
----- Project Information
Center project name: L3324
Center clone name: 13_G_16

* NOTE: This is a 'working draft' sequence. It currently
* consists of 55 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.

1 1382: contig of 1382 bp in length
* 1383 1482: gap of 100 bp
* 1483 2509: contig of 1027 bp in length
* 2510 2609: gap of 100 bp
* 2610 3615: contig of 1006 bp in length
* 3616 3715: gap of 100 bp
* 3716 5107: contig of 1392 bp in length
* 5108 5207: gap of 100 bp
* 5208 6331: contig of 1124 bp in length
* 6332 6431: gap of 100 bp
* 6432 7570: contig of 1139 bp in length
* 7571 7670: gap of 100 bp
* 7671 8689: contig of 1019 bp in length
* 8690 8789: gap of 100 bp
* 8790 9889: contig of 1100 bp in length
* 9890 9989: gap of 100 bp
* 9990 11351: contig of 1362 bp in length
* 11352 11451: gap of 100 bp
* 11452 13214: contig of 1763 bp in length
* 13215 13314: gap of 100 bp
* 13315 14472: contig of 1158 bp in length
* 14473 14572: gap of 100 bp
* 14573 16043: contig of 1471 bp in length
* 16044 16143: gap of 100 bp
* 16144 17521: contig of 1378 bp in length
* 17522 17621: gap of 100 bp
* 17622 19059: contig of 1438 bp in length
* 19060 19159: gap of 100 bp
* 19160 20596: contig of 1437 bp in length
* 20597 20696: gap of 100 bp
* 20697 21997: contig of 1301 bp in length
* 21998 22097: gap of 100 bp
* 22098 24537: contig of 2440 bp in length
* 24538 24637: gap of 100 bp
* 24638 25795: contig of 1158 bp in length
* 25796 25895: gap of 100 bp
* 25896 27194: contig of 1299 bp in length
* 27195 27294: gap of 100 bp
* 27295 29425: contig of 2131 bp in length
* 29426 29525: gap of 100 bp
* 29526 31101: contig of 1576 bp in length
* 31102 31201: gap of 100 bp
* 31202 33081: contig of 1880 bp in length
* 33082 33181: gap of 100 bp
* 33182 34843: contig of 1662 bp in length
* 34844 34943: gap of 100 bp
* 34944 37369: contig of 2426 bp in length
* 37370 37469: gap of 100 bp
* 37470 39552: contig of 2083 bp in length
* 39553 39652: gap of 100 bp
* 39653 42220: contig of 2568 bp in length
* 42221 42320: gap of 100 bp
* 42321 44105: contig of 1785 bp in length
* 44106 44205: gap of 100 bp
* 44206 45952: contig of 1747 bp in length
* 45953 4715: contig of 1563 bp in length
* 4716 47715: gap of 100 bp
* 47716 51120: contig of 3405 bp in length
* 51121 51220: gap of 100 bp
* 51221 53770: contig of 2550 bp in length
* 53771 53870: gap of 100 bp
* 53871 56963: contig of 3093 bp in length
* 56964 57063: gap of 100 bp
* 57064 59799: contig of 2736 bp in length
* 59800 59899: gap of 100 bp
* 59900 63301: contig of 3402 bp in length
* 63302 63401: gap of 100 bp
* 63402 64177: contig of 776 bp in length
* 64178 64277: gap of 100 bp
* 64278 69030: contig of 4753 bp in length
* 69031 69130: gap of 100 bp
* 69131 73519: contig of 4389 bp in length
* 73520 73619: gap of 100 bp
* 73620 76256: contig of 2637 bp in length
* 76257 76356: gap of 100 bp
* 76357 80598: contig of 4242 bp in length
* 80599 84156: contig of 3458 bp in length
* 84157 84256: gap of 100 bp
* 84257 89113: contig of 4857 bp in length
* 89114 89213: gap of 100 bp
* 89214 93633: contig of 4420 bp in length
* 93634 93733: gap of 100 bp
* 93734 98328: contig of 4595 bp in length
* 98329 98428: gap of 100 bp
* 98429 102945: contig of 4517 bp in length
* 102946 103045: gap of 100 bp
* 103046 108095: contig of 5050 bp in length
* 108096 108195: gap of 100 bp
* 108196 114837: contig of 6642 bp in length
* 114838 114937: gap of 100 bp
* 114938 121563: contig of 6626 bp in length
* 121564 121664: gap of 100 bp
* 121665 128405: contig of 6742 bp in length
* 128406 128505: gap of 100 bp
* 128506 133978: contig of 5473 bp in length
* 133979 134078: gap of 100 bp
* 134079 138994: contig of 4916 bp in length
* 138995 139094: gap of 100 bp
* 139095 146606: contig of 7512 bp in length
* 146607 146706: gap of 100 bp


```
/gene="5-HT5B"
/pseudo
/codon_start=1
/product="serotonin receptor"
/db_xref="PSEUDO:CAC36299.1"
/db_xref="REMTREMBL:CAC36299"
<649...907
/gene="5-HT5B"
/number=1
/pseudo

exon

ORIGIN
Query Match 64.7%; Score 745; DB 9; Length 977;
Best Local Similarity 94.6%; Pred. No. 1.9e-114;
Matches 803; Conservative 0; Mismatches 20; Indels 26; Gaps 2;

QY 1 CGCMTAGAGCGCGTAGCTTTAGTGGCGACCGCGCGGTGCGCTTGCCTGGGACC 60
DB 85 GCGCGTGGAGCGCGTAGCTTTAGTGGCGACCGCGCGGTGCGCTTGCCTGGGACC 144
QY 61 CGAGACACAGCAGC-GGACCCGGGACCCCAAGCCCGAGAGGATATCTCGGTTCCGACCCGA 119
DB 145 CGAGACACAGCAGCAGGACCCGGGACCCCAAGCCCGAGAGGATATCTCGGTTCCGACCCGA 204
QY 120 GCGCGCGCGTCTGCGCGGCGGAGGCGCGCTTCTCTGTCTTCAAGTCTGTGGTGA 179
DB 205 GCGCGCGCGTCTGCGCGGCGGAGGCGCGCTTCTCTGTCTTCAAGTCTGTGGTGA 264
QY 180 CGCTGCTAGTGTGTGTATGCTGTGCGACCTTCTGCTGTGGAACCTGCTGCTGCGTCA 239
DB 265 CGCTGCTAGTGTGTGTATGCTGTGCGACCTTCTGCTGTGGAACCTGCTGCTGCGTCA 324
QY 240 TCCCGCGGTGCGTGTCTTCCACCGCGTGGCGGATTAATCTGTGTGGCTGCGGCGGTCT 299
DB 325 TCCCGCGGTGCGTGTCTTCCACCGGTGGCGGATTAATCTGTGTGGCTGCGGCGGTCT 384
QY 300 CGGACGAACTAGTGGCAGCGCTGGCGATGCCACCGAGCGTGGCGAGTGTGACCG 359
DB 385 CGGACGAACTAGTGGCAGCGCTGGCGATGCCACCGAGCGTGGCGAGTGTGACCG 444
QY 360 GCGCAGTGGTGTGTGGCGCGGAGCTGTGCGAGTGTGATCTCTTCCGCGCGGAG 419
DB 445 GCGCAGTGGTGTGTGGCGCGGAGCTGTGCGAGTGTGATCTCTTCCGCGCGGAG 504
QY 420 CTGTG-----CTGTGTGCGCGCGCGCGCTCGGGAACGTG 454
DB 505 CTGTGCGACGTGTGATCTCTTCCACGCGTGTGCTGCGCGCGCGCTCGGGAACGTG 564
QY 455 GCGGCGATCGCGTGGCGCGGACCGGCGCATACACGCGCATGCGACACAGTGTGCG 514
DB 565 GCGGCGATCGCGTGGCGCGGACCGGCGCATACACGCGCATGCGACACAGTGTGCG 624
QY 515 ACCCGAGCGCGCTGCTGTGCTATGATCGGCTGCGCGCGGTCGCGTGGCGCTCATC 574
DB 625 ACCCGAGCGCGCGCTGCTGTGCTATGATCGGCTGCGCGCGGTCGCGTGGCGCTCATC 684
QY 575 GCGCTGCGCGCTGCTGCTTTGGCGCGGCGAGTGTGCGAGCTGCGCTCCAGCGCTGC 634
DB 685 GCGCTGCGCGCTGCTGCTTTGGCGCGGCGAGTGTGCGAGCTGCGCTCCAGCGCTGC 744
QY 635 CAGGTGAGCGCGGAAACCTCTATGCGCGCTTCTCCACCGCGGCGCGCTTCCACCTGCG 694
DB 745 CAGGTGAGCGCGGAAACCTCTATGCGCGCTTCTCCACCGCGGCGCGCTTCCACCTGCG 804
QY 695 CTTGGGTGGTGGCTGCTGTCTACCGGAAGATCTAGAGGCGCGCAAGTTTGTGTTTCGG 754
DB 805 CTTGGGTGGTGGCTGCTGTCTACCGGAAGATCTAGAGGCGCGCAAGTTTGTGTTTCGG 864
QY 755 CGCGCGCGGAGAGCTGTGCTGCGCTTGGCGGCGCACCATGCGAGTGAAGTCCAAGTAAAG 814
DB 865 CGACGCGGAGAGAGCTGTGCTGCGCTTGGCGGCGCACCATGCGAGTGAAGTGGGCTGAG 924
QY 815 GAAGCACT 823
```

```

666 CGTTGGCAGCTAGGACGAGTCTGTGCCAGGTGTGATCTCTTCGACG-----TG 716
425 CTGTGTCCTCCCGCGGCTCGGAGACGTGGCGGCATCGCCCTGGCGCGGACGCGGCC 484
717 TTGTGCTGCACCGCAGCATCTGGAACGTGGCGGCATCGCCCTGGATCGCTACTGGACT 776
485 ATCACAGGCACCTGACAGCATCGTGGCGACCCGCGAGCGCGCTGTGCTCATGATC 544
777 ATCAGCGCACCTGACAGTACAGCTGCGACCCCGAGCGCGTCTCTGCGCTCATGATC 836
545 GCGTGTGCGCGGTCGCTCGGCGCTCATCGCCCTGCGCGCGCTCTCTTTGGCGGCGGC 604
837 GCGATCACCTGGGCACTCTCGCGCTCATGCTCTCGCCCGCTGCTTTTGGCTGGGCG 896
605 GAGTGTGCGAGCTCGGCTCCAGGCTCGGAGTGCAGCGTGCAGCGGGAACCTCTCATGCCGCC 664
897 GAAGCCTATGATGTCTGGCTCGAGCGTTGCCAGGTGAGCCAGGAGCCCTCTATGCTGTC 956
665 TTCTCCACCCCGCGGCTCTCCACTGCGGCTTGGCGTGTGCGGTTTGTCTACCGGAAG 724
957 TTCTCCACCTCGGAGCCTTCTACTGCTCTAGCGGTGGTGTCTCTGCTCTACTGGAAA 1016
725 ATCTACGAGCGCGCAAGTTTCGTTTCGCGCGCGCGCGAGAGCTGTGCTGCGGTTGCCG 784
1017 ATATACAAAGCGCCAAAGTTTCGATTGCTGCGAGCGCGGCGGTGTGATCCGCTTCT 1076
785 GCCACCATGAGGTGAGTCCAGGTAAAGAACACCTGATGAGCTGAAGTGTGTTTC 844
1077 GCCACCAAGC-----AGCAAAGGAAGCACCTCGGAGTCTGAGATGTGTTTC 1124
845 ACGCACATTCAAAGCAAGGTGTCTTCAGGTGAGCGGGACTCTCGCGGGAGCAG 904
1125 ACAGCCGTCGCCAGCAGTACGTGACCTTCAGACAGCGAGAGCTCTCGCGGAGCAG 1184
905 AAGGAGGCGGAGCAGCATGATGTTGGGAATTCGATTGCGGTGTTTGTGCTGTGCG 964
1185 AAGGAGAAGCGGCGAGCATGATGTTGGGATTCGATTGCGGTGTTTGTGCTGTGCG 1244
965 ATCCCTCTCTCTGAGGAACTATAGCCACTCTGTGCTGTGAGCTGCGCCGCCATC 1024
1245 ATCCCTCTCTCTGAGGAGCTATCAGCCCGCTCTGTGCTGTGAGCTGCGCCGCCATC 1304
1025 TGGAAAGCATATTCGTGCTGCTGCTTCTTCAATTCCTTCAACCCCTGATTTAC 1084
1305 TGGAAAGCATATTCGTGCTGCTGATATTCATTCGTTTCAACCCCTGATTTAC 1364
1085 ACAGCTTTTAAAGAACTACAAATGCTTCAAGAGCCTCTTTTACTAAGCAGATGA 1144
1365 ACTGCTTTTAAAGAACTACAAATGCTTCAAGAGCCTCTTTTACTAAGCAGATGA 1424

RESULT 13
MM5HT5BSR 2061 bp mRNA linear ROD 26-MAY-1993
LOCUS M.musculus mRNA encoding 5-HT5B serotonin receptor.
DEFINITION X69867
ACCESSION X69867
VERSION X69867.1 GI:288735
KEYWORDS 5-HT5B serotonin receptor.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 2061)
Matthes,H., Boschert,U., Amlaiky,N., Grailhe,R., Plassat,J.L.,
Muscatelli,F., Mattei,M.G. and Hen,R.
Mouse 5-hydroxytryptamine5A and 5-hydroxytryptamine5B receptors
define a new family of serotonin receptors: Cloning, functional
expression, and chromosomal localization
Mol. Pharmacol. 43 (3), 313-319 (1993)
93196607
MEDLINE PUBMED 8450829
REFERENCE 2 (bases 1 to 2061)

```

```

AUTHORS Matthes,H.
Direct Submission
TITLE Submitted (28-DEC-1992) H. Matthes, Laboratoire de Genetique
JOURNAL Moleculaire des, Eucaryotes, Departement de Neurobiologie, 11 rue
Humann, F-67000 Strasbourg, FRANCE
FEATURES
source
1. .2061
/organism="Mus musculus"
/mol_type="mRNA"
/db_xref="taxon:10090"
/tissue_type="brain"
15. .2055
15. .311
312. .1424
/codon_start=1
/product="5-HT5B serotonin receptor"
/protein_id="CAA49501.1"
/db_xref="GI:288736"
/db_xref="GOA:P31387"
/db_xref="SWISS-PROT:P31387"
/translation="MEVSNLSGATPLAFPPGPESCDSPSSGRSMGTPGGLILPGR
EPESAPTVLVVLLVLLIATFLNLLVLLVILVRFAHRVPHNLVASTAVSVLVA
VLVNPGLSVLSLSAGRWQLSRSLCHVWISFDVLCCTASINWAAIALDRWTITRHL
QYTLRTSRASALMIATLWALSIALIPLLFGWEAYDARLQRCQVQSESVAVFST
CGAFYLRPLAVLVFVWKIYKAAKPRFGRRRRAVPLPATCOAKEAPSEMVFTARR
ATVTFQTSVGSWSREQEKRAMVMVGLILGVFLCNIPFFLITELISPLCACSLLPIWKS
IFLWLGYSNFFNELIYIAFNKNNNNAKSLFTIKQR"
312. .1088
/label=exon1
1089. .1421
/label=exon2
1425. .2055
exon
exon
5'UTR
ORIGIN
Query Match 63.6%; Score 732.8; DB 10; Length 2061;
Best Local Similarity 79.9%; Pred. No. 1.8e-112;
Matches 911; Conservative 0; Mismatches 202; Indels 27; Gaps 3;
QY 5 ATGAGGCGGCTAGCCTTTAGTGGCCACCGCCGCGGCTTGCCCTTGCCCTGGGACCGGAG 64
DB 312 ATGGAAGTTTCTAACTCTCAGGCGCCACTCCCGCGCTTGCTTCTCCGCGGACCTGAG 371
QY 65 ACCAGCAGCGGACCGGAGCCCAAGCCCGAGAGGATCTCGGTTCCAGCCCGGAGCGGC 124
DB 372 AGCTGCAGTGAC-----AGCCCAAGTTCGCGAGAGCATGGGATCCACCCAGGTGG 425
QY 125 GCCGTCTCGCGGCGGAGGCGCCCTTCTCTCTTTCACGCTCTGTTGGTGCAGCGTG 184
DB 426 CTCATCTTGGCGGCGGAGCGCCCTTCTCTGCTTTCACCGTGTGTGTGACTCTA 485
QY 185 CTAGTGTGTGATGCTGCGCCACTTCTGTGGAACCTGCTGGTTCGGTCCACCATCCG 244
DB 486 CTGTTGTGTGATGCTGCGCCACTTCTTATGGAATCTGTAGTCTGTGTAATCTGCT 545
QY 245 CGGCTCGTGTGCTTCCACCGGCTCCGCATAACTTGTGGCTCGAGCGGCTCTCGGAC 304
DB 546 CGGCTCGGCGCTTCCACCGGCTCCACATAACTTGTGGCTCGAGAGCGCTCTCGAT 605
QY 305 GAAGTGTGCGGCTGGCGATGCCACCGAGCCTGGCGAGTGTGAGCTGTGACCGGCGGA 364
DB 606 GTCCTGTGCGGCTTCTGTGTGATGCTGCTGAGCCCTGAGCGAGTGTGCGTGGGCGA 665
QY 365 CGTGGCTGTGGCGGAGCGCTGTGCGAGTGTGATCTCTCTTCGAGCGCGGAGCGCTGT 424
DB 666 CGTTGGCAGCTAGGACGAGTCTGTGCGAGCTGTGAGTCTCTCTTCGACG-----TG 716
QY 425 CTGTGCTGCCCGCGGCTCGGGAACGTGCGGCCATCGCCCTGGCGCGGCGGCGGCC 484
DB 717 TTGTGCTGCAACCGCAGCATCTGGAACGTGGCGCCATCGCCCTGGATCGCTACTGGACT 776
485 ATCACAGGCACCTGACAGCATCGTGGCGACCCGCGAGCGCGCTCTGTTGCTCATGATC 544
777 ATCACGCGCACCTGACAGTACAGCTGCGACCCCGAGCGCGTCTCTGCGCTCATGATC 836

```

```
QY 545 GCGCTCGCCGCGGTCGCGTGGCGCTCATCGCCCTCGCGCGCTGCTCTTTGGCCGGGCG 604
Db 837 GCGATCACTGGGCACTGTCGCGCTCATGTCTCGCCCGCTGCTTTTGGCTGGGCG 896
QY 605 GAGGTGTGCGAGCTCGGCTCCAGCGCTCCAGGTGAGCCGGAAACCCCTCTATGCGGCC 664
Db 897 GAAGCTATGATGCTCGGCTGAGCGTTGCCAGGTGAGCCAGGAGCCCTCTATGCTGTC 956
QY 665 TTCTCAACCGGGGCGCTTCACCTGCGGCTTGGGTGGTGGCTTGTCTACCGGAAG 724
Db 957 TTCTCAACCTGGGAGCGCTTACCTGCGCTTAGCGGTGGTGTCTCTACTGGAAA 1016
QY 725 ATCTAGAGCGGCGCAAGTTTCGTTTCGCGCGCGCGGAGAGCTGTGTGCTGCGTTCGCG 784
Db 1017 ATATCAAAAGCCGCAAGTTTCGATTCGTTGCGAGACGGCGGGGTGGTACCGCTTCCT 1076
QY 785 GCCACCATGAGGTGAGGTCCAGGTAAAGGAAGCACCTGATGAGGCTCAAGTGTGTTTC 844
Db 1077 GCCACCACGC-----AGGCAAGGAAGCACCTCCGAGTCTGAGATGGTGTTC 1124
QY 845 AGCGACATTTGAAAGCAAGGTGCTCTTCCAGGTGAGCGGGGACTCCTGGCGGGAGCAG 904
Db 1125 ACAGCCCGTTCGCGGACNACAGTGACCTTCCAGACAAAGCGGAGACTCCTGGCGGGAGCAG 1184
QY 905 AAGGAGAGCGGAGCAGCAGATGATGGTGGAAATTCGATTTGGCGTGTTCGTGTGCTGG 964
Db 1185 AAGGAGAGCGGAGCAGCAGATGATGGTGGGATCTTGATTTGGCGTGTTCGTGTGCTGG 1244
QY 965 ATCCCTCTTCTCTGACGGAACATCATCAGCCCACTCTGTGCTGAGCCTGCCCGCCATC 1024
Db 1245 ATCCCTCTTCTCTGACGAGCTCATCAGCCGCTCTGTGCTGAGCCTGCCACCATC 1304
QY 1025 TGGAAAGCATATTTCTGTGGCTTGGCTACTCAATTTCTTCAACCCCTGATTTAC 1084
Db 1305 TGGAAAGCATATTTCTGTGGCTTGGATATTCATTTCTTCAACCCCTGATTTAC 1364
QY 1085 ACAGCTTTTAAAGAACTACAAATGCTTCAAGAGCTCTTTACTAAGCAGAGATGA 1144
Db 1365 ACTGCCTTTAAAGAAATACAAATGCTTCAAGAGCTCTTTACTAAGCAGAGATGA 1424

RESULT 14
AR080627
LOCUS AR080627 2226 bp DNA linear PAT 31-AUG-2000
DEFINITION Sequence 1 from patent US 5968817.
ACCESSION AR080627
VERSION AR080627.1 GI:10007357
KEYWORDS
SOURCE Unknown.
ORGANISM Unclassified.
REFERENCE
1 (bases 1 to 2226)
AUTHORS Sutcliffe,J.Gregor., Erlander,M.G. and Lovenberg,T.W.
TITLE DNA encoding serotonin receptors
JOURNAL Patent: US 5968817-A 1 19-OCT-1999;
FEATURES
Location/Qualifiers
source
1..2226
/organism="unknown"
/mol_type="unassigned DNA"

ORIGIN
Query Match 63.4%; Score 729.8; DB 6; Length 2226;
Best Local Similarity 79.6%; Pred. No. 5.5e-112;
Matches 911; Conservative 0; Mismatches 207; Indels 27; Gaps 3;

QY 3 CCATGGAGGCGGTAGCTTTCACTTCACTGAGTGGCCACCGCGGCTTGCCTTGGCCCTGGAGCCCG 62
Db 301 CAATGGAAGTCTTAACCTCTCAGGCGCCACCGCTTGGCATTTCTCTCCGGGACCCG 360
QY 63 AGACAGCAGCGGACCGGACCCCGACCCGAGAGGGATACTCGTTCGACCCCGAGCG 122
Db 361 AGAGCTCAGTGAC-----AGCCCAAGTTCCGCGCAGAGCATGGGATCCACCCCGAGGG 414
```

```
QY 123 GCGCGTCTCTGCGGCGCGAGGCGCGCCCTTCTCTGTCTTTCACGGTCTCTGGTGGTACGC 182
Db 415 GACTCATCTTGTTCGGGCGCGAGCGCCCTTCTCTGCCCTTACCGTACTCGTGGTAACTC 474
QY 183 TCTAGTGTCTGTATTCGCTGCCACTTTCCTGTGAACTCTGTGTGGTTCGGTCCGATCATCC 242
Db 475 TACTGTGTGTGATCGCTGCCACTTTCCTATGAACTCTGTAGTCTTCTGGTGAATATCC 534
QY 243 CGCGGTCGCTGCTTCCACCGGTCGCGATAACTTGTGTGCTCTGACGGCGCTCTCG 302
Db 535 TCGCGCTCGCGCCCTTCCACCGTGTGCCACATAACTTGTGTAGCTCTGACCGCGCTCTCG 594
QY 303 AGCAACTAGTGTGAGCGCTTGGCGATGCGCACCGAGCTTGGCGAGTGTGTACCGGCG 362
Db 595 AGCTCTGTGTGGGCTCTGTGTGATGCCACTGAGCTGTGTGAGCGAGTGTGTGGCTGGCG 654
QY 363 GACGTGCTGTGTGGCGCGAGCGCTGTGCGACGTGTGTGATCTCTCTGACGCGCGGAGCCT 422
Db 655 GACGTGTGCGAGCTGGCGAGGAGTCTGTGCGACGTGTGTGATCTCTCTGACG----- 705
QY 423 GTCTGTGTGCTGCGCGCGCTCGGGAAGCTGCGGCCCATCGCCCTGGCGCGCGAGCGGG 482
Db 706 TGTGTGTGTGCGACAGCCAGCATCTGGAACGTGTGGGCCCATGCGCCCTGGATCGCTACTGGA 765
QY 483 CCATCACAGCGCACCTTGCAGCACAGCTCGCACCCCGCAGCCCGCTCTCTGTCTCATGA 542
Db 766 CTATCACGCGCCACCTTGCAGTACACGCTCGCACCCCGCGCGCTTTCAGCACTTATGA 825
QY 543 TCGCGTCTCGCGGTCGCGTGTGCGCGCTCATGCGCTCGCGCGCTGCTCTTGGCGCGG 602
Db 826 TCGCGATCACTTGGCGACTGTGCGCGCTCATGCGCTCGCGCGCTGCTCTTGGCGTGG 885
QY 603 GCGAGTGTGCGACGCTCGGCTTCCAGCTGCGAGTGAGCGGGAACCTCTCTATGCGG 662
Db 886 GCGAAGCTTACGATGCTCGGCTGAGCGTTCAGGTGAGCCAGGAGCTTCTGTAGCGCG 945
QY 663 CTTCTTCCACCGCGCGCTTCCACCTCGCGCTTGGCGTGTGCGGCTTGTCTACCGGA 722
Db 946 TCTTCTCCACCTGCGGAGCTTCTAGTGCCTTGGCGCTGCTCTTCTGCTACTGGA 1005
QY 723 AGATCTAGAGGGGCGAAGTTCGTTTGGCGCGCGCGGAGAGCTGTGCTGCGCTGC 782
Db 1006 AGATATAAAGCGCGCAAGTTTCGATTTCGCGCGCAGACGCGGCGGCTAGTCCCGCTC 1065
QY 783 CGGCAACCATGAGGTGAGGTCCAGGTAAAGGAACACTGATGAGGCTGAAGTGTGT 842
Db 1066 CGGCCACACGC-----AGGCAAGGAAGCACTCTAGGAGTCTGAGACGGTAT 1113
QY 843 TCACGGCACATTCGAAAGCAAGGTGTCTCTCCAGGTGAGCGGGAACCTCTGCGCGGAGC 902
Db 1114 TCACCGCGCTTTCAGAGCGACAGTGCCTTCCAGACAAGTGGAGACTCTGCGCGGAGC 1173
QY 903 AGAAGAGAGCGGAGCAGCAGCATGATGCTGGGAATCTGATGCGGTGTTGTGCTGTGCT 962
Db 1174 AGAAGAGAGCGGAGCAGCAGCATGATGCTGGGAATCTGATGCGGTGTTGTGCTGTGCT 1233
QY 963 GATATCCCTTCTTCTGACGGAACCTCATCAGCCCACTCTGTGCTGCGAGCTGCGCCCGCA 1022
Db 1234 GATATCCCTTCTTCTGACGAGCTCTGTCAGCCCGCTCTGCGCTGCGAGCTGCGCACCGCA 1293
QY 1023 TCTGAAAAGCATATTTCTGTGCTTGGCTTACTTCCAAATTTCTTTTCAACCCCGCTGATTT 1082
Db 1294 TCTGAAAAGCATATTTCTGTGCTTGGCTTACTTCAAAATTCGTTCTTCAATCCCTTAACT 1353
QY 1083 ACACAGCTTTTAAAGAACTACAACAATGCCCTTCAAGGCTCTTTTACTTAAGCAGAT 1142
Db 1354 ACACGCTTTTAAAGAACTACAACAATGCCCTTCAAGGCTCTTTTACTTAAGCAGAT 1413
QY 1143 GAACA 1147
Db 1414 AAGCA 1418
```

```

RESULT 15
RAT5HT5B          2240 bp      mRNA      linear      ROD 26-JUL-1993
LOCUS              Rattus norvegicus 5-hydroxytryptamine receptor (5HT5b) mRNA, 5'
DEFINITION
ACCESSION          L10073
VERSION            L10073.1 GI:310074
KEYWORDS            5-hydroxytryptamine receptor; serotonin receptor.
SOURCE              Rattus norvegicus (Norway rat)
ORGANISM            Rattus norvegicus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.
REFERENCE           1 (bases 1 to 2240)
AUTHORS             Erlanger,M.G., Lovenberg,T.W., Baron,B.M., Lecca,L.,
                    Danielson,P.E., Rache,M., Stone,A.L., Siegel,B.W., Foye,P.E.,
                    Cannon,K., Burns,J.E. and Sutcliffe,G.J.
                    Two members of a distinct subfamily of 5-hydroxytryptamine
                    receptors differentially expressed in rat brain
JOURNAL             Proc. Natl. Acad. Sci. U.S.A. 90 (8), 3452-3456 (1993)
MEDLINE             93234515
PUBMED              7682702
COMMENT              Original source text: Rattus norvegicus (strain Sprague-Dawley)
                    hypothalamus cDNA to mRNA.
FEATURES             Location/Qualifiers
                     1..2240
                        /organism="Rattus norvegicus"
                        /mol_type="mRNA"
                        /strain="Sprague-Dawley"
                        /db_xref="taxon:10116"
                        /tissue_type="hypothalamus"
                     303..1415
                        /function="unknown effector coupling"
                        /standard_name="5-hydroxytryptamine receptor"
                        /note="putative"
                        /codon_start=1
                        /label="5-HT5b"
                        /product="serotonin receptor"
                        /protein_id="AAA40616.1"
                        /db_xref="GI:310075"
                        /translation="MEVSNLSGATPGIAPFPSPSCDPSGRSMGSTPGILLISGR
EPPSFTLVVTLVLLIAFTWNLVLVTLIRVAFRPHVPHNLVASTVSDVIVA
ALVPLSLVLSLSGRWOLGRSLCHVWISFDVLCCTASINWVAIALDRVYTLTRHL
QYTLRRASALMIATWALSIALIALPFLPGMEAVDAELQRCVQSOEPSYAVFST
CGAVPLAVLVYVYKLYKAKRFGRERRRNVPLPATYQAKAPQESRVPFARCR
ATVAFQISGSWRQKRAAMVIGLIGFVLWCWIPFLTELVSPLCACSLPPIPKWS
IFLWIGYSNFNPLLYTAFNKNYNNAPKSLFTKQR"
ORIGIN
Query Match          63.4%; Score 729.8; DB 10; Length 2240;
Best Local Similarity 79.6%; Pred. No. 5.5e-112;
Matches 911; Conservative 0; Mismatches 207; Indels 27; Gaps 3;

QY      3 CCATGGAGCGCGTAGCCTTTTCAGTGGCCACCGCGCGGTGGCCCTTGCCTTGGACCCG 62
DB      301 CAATGGAAGTCTCTAACTCTCAGCGCCACCCCTGGCATTGCCCTTCTCCTCGGACCCG 360

QY      63 AGACGAGGAGCGGACCGGACCCAGCCGAGAGGATACCTCGGTCGACCCCGAGCG 122
DB      361 AGAGTCGAGTGAC-----AGCCCAAGTTCGCGCAGAAGCATGGGATCCACCCGAGGG 414

QY      123 CGCGCGTCTCGCGGCGGAGCGCGCCCTTCTCTGCTTCCACGCTCTCGTGGTGAGCG 182
DB      415 GGCTCAUCTTGTCCGCGCGGAGCGCGCCCTTCTCTGCTTCCACGCTCTCGTGGTAACTC 474

QY      183 TGTAGTGTCTGATCGCTGCCAATTTCTCTGTGGAACCTCTGTTCCGGTCCACCATCC 242
DB      475 TACTGGTGTCTGATCGCTGCCACTTTCTTATGGAATCTCTAGTCTGTGTGACTATCC 534

QY      243 CGCGGTCGCGCTTCCACCGGTGCGCGAATCTTGGTGGCCTCGACGCGCGGTCTCGG 302
DB      535 TSCGCGTCCGCGCTTCCACCGGTGCGCACATACTTGGTAGCCTCGACGCGCGGTCTCGG 594

```

```

QY      303 ACRAACTAGTGGCAGCGCTGGCGATGCCACGAGCCTGCGAGTGAGCTGTCTGACCGGCG 362
DB      595 ACGTCTGTGGCGGCTCTGGTGATGCCACGTAGCCTGTGAGCGAGTTGTCTGGCTGGCG 654

QY      363 GACGTGGCTGTCTGGCGCGGAGCCTGTGTCCACGTGTGGATCTCTTTCGACGCGGAGCCT 422
DB      655 GACGTTGGCAGCTGGCGCAGGAGTCTGTGCCACGTGTGATCTCTCTTCGACG----- 705

QY      423 GTCTGTGTCTGCCCGCGCGCCTCGGGAACGTGTGGCGGCATCGCTTGGGCGCGGACGGG 482
DB      706 TGTGTGTGTGCACACGCCAGCATCTGGAACGTGTGGCGGCATCGCTTGGATCGCTACTGGA 765

QY      483 CCATCACACGCGACCTGTCAGCACACGCTGCCACCGCGCAGCGCGCTCGTGTCTCATGA 542
DB      766 CTATCACCGGCACCTGTCAGTACACGCTGGCACCCGCGCGCGCTTCAGACATTATGA 825

QY      543 TCGCGCTCGCGCGGTGGCTGGCGCTCATCGCCCTCGCGCGCGCTGTCTTTGGCGCGGG 602
DB      826 TCGGATCACCTGGGCACTGTCCGCCCTCATCGCGCTCGCCCGCTGTCTTTGGCTGGG 885

QY      603 GCGAGGTGTGCAGCTCGGCTTCAGCGCTCGCAGGTGAGCGCGGGAACCTCTCTATGCGG 662
DB      886 GCGAAGCCTACGATGCTCGGCTCGAGCGTTGCCAGGTGAGCGCAGGAGCTTCGTACGCGG 945

QY      663 CTTTCTCCACCGCGCGCGCTTCCACCTGCGCGCTTGGCGTGGTGGCTTGTCTACCGGA 722
DB      946 TCTTCTCCACCTGCGGAGCTTTTACGTGCTCTGGCGGTGGTGTCTTCTGTCTACTGGA 1005

QY      723 AGATCTACGAGCGCGCAAGTTTCGTTTCGCGCGCGCGCGGAGAGCTGTGTGCGGTGCG 782
DB      1006 AGATATACAAAGCGCCCAAGTTTCGATTTCGCGCGCGCAGACGCGCGGTAGTGGCCCTGC 1065

QY      783 CGGCAACCATCGAGTGAAGTCCAAAGTAAAGGAAGCACCTGTATGAGGCTGAAGTGGTGT 842
DB      1066 CCGCCACCCACGC-----AGGCAAGGAAGCACCTCAGGAGTCTGAGACGCTAT 1113

QY      843 TCAGGCGACATTGCAAGAACGCTGTCTTCCAGGTGAGCGGAGACTCCTTGGCGGAGC 902
DB      1114 TCACGCGCGTTGCAGAGCGACAGTGGCTTTCACAGCAAGTGGAGACTCCTTGGCGGAGC 1173

QY      903 AGAAGGAGGAGCGAGCAGCCATGATGTGGGAATTTCTGATTGGCGTGTGTTGTGTGTGCT 962
DB      1174 AGAAGGAGAGCGAGCGCCCATGATGTGGGGAATCTTGATCGGTGTGTTGTGTGTGCT 1233

QY      963 GGATCCCGCTTCTCTGACGGAACATCATACGCCACTCTGTGCGCTGACGCTGCCGCCCA 1022
DB      1234 GGATCCCGCTTCTCTGACGAGCTCGTACGCCGCTCTGCGCGCTGCGAGCCTTGCACCCA 1293

QY      1023 TCTGGAAGAGCATATTTCTGTGGCTTGGCTACTCCAAATTTCTTCAACCCCTGATTT 1082
DB      1294 TCTGGAAGAGCATATTTCTGTGGCTTGGCTATTCAAAATTCGTTCTTCAATCCCTTAATCT 1353

QY      1083 ACACAGCTTTTAAACAAGAACTACAACTATGCTTTCAAGAGCCTCTTTACTAAGCAGAGAT 1142
DB      1354 ACACGCGCTTTTAAAGAAGAACTACAACTATGCTTTCAAGAGCCTCTTTACTAAGCAGAGAT 1413

QY      1143 GAACA 1147
DB      1414 AAGCA 1418

```

Search completed: September 7, 2004, 16:25:20
Job time : 4463 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - protein search, using frame_plus_n2p model

Run on: August 31, 2004, 20:09:52 ; Search time 30 Seconds
(without alignments)
3964.875 Million cell updates/sec

Title: US-09-976-782-15

Perfect score: 2154

Sequence: 1 cgcctaggagccgtagcc.....aagcagagatgaacacaggg 1152

Scoring table:

BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 399414 seqs, 51625971 residues

Total number of hits satisfying chosen parameters: 778828

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+n2p.model -DEV=xl
-Q=/cgn2_1/USPTO_spool_p/US09976782/runat_31082004_115952_10354/app_query.fasta_1.1351
-DB=Issued Patents AA -QMT=fastan -SUFFIX=n2p.ra1 -MINMATCH=0.1 -LOOPCL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blomsum62 -TRANS=human40.cdi
-LIST=45 -DOALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=ptc -NORW=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09976782@cgn1_1_46@runat_31082004_115952_10354 -NCPU=6 -ICPU=3
-NO_MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

Issued Patents AA:*
1: /cgn2_6/ptodata/2/iaa/5A.COMB.pep.*
2: /cgn2_6/ptodata/2/iaa/5B.COMB.pep.*
3: /cgn2_6/ptodata/2/iaa/6A.COMB.pep.*
4: /cgn2_6/ptodata/2/iaa/6B.COMB.pep.*
5: /cgn2_6/ptodata/2/iaa/PCTUS.COMB.pep.*
6: /cgn2_6/ptodata/2/iaa/backfiles.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1444.5	67.1	370	2	US-08-031-538-2
2	1022.5	47.5	357	1	US-08-356-405-2
3	1022.5	47.5	357	1	US-08-356-405-9
4	1014.5	47.1	357	2	US-08-031-538-4
5	448	20.8	390	2	US-08-461-812-4
6	448	20.8	390	2	US-08-157-185-15
7	448	20.8	390	3	US-08-281-526B-15
8	448	20.8	390	4	US-09-450-790A-15
9	448	20.8	390	4	US-09-332-837-15
10	448	20.8	390	4	US-09-371-705-4
11	433	20.1	390	1	US-07-817-920-6
12	433	20.1	390	1	US-08-117-006-6

13	433	20.1	390	1	US-08-216-594-6	Sequence 6, Appli
14	433	20.1	390	5	PCT-US93-00149-6	Sequence 6, Appli
15	433	20.1	398	1	US-08-370-542-6	Sequence 6, Appli
16	433	20.1	398	1	US-08-542-358-6	Sequence 6, Appli
17	433	20.1	398	3	US-09-018-351-6	Sequence 6, Appli
18	431.5	20.0	375	1	US-08-370-542-5	Sequence 5, Appli
19	431.5	20.0	375	1	US-08-542-358-5	Sequence 5, Appli
20	431.5	20.0	375	3	US-09-018-351-5	Sequence 5, Appli
21	431.5	20.0	376	1	US-07-817-920-5	Sequence 5, Appli
22	431.5	20.0	376	1	US-08-117-006-5	Sequence 5, Appli
23	431.5	20.0	376	1	US-08-216-594-5	Sequence 5, Appli
24	431.5	20.0	376	5	PCT-US93-00149-5	Sequence 5, Appli
25	430	20.0	377	2	US-08-461-812-2	Sequence 2, Appli
26	430	20.0	377	2	US-08-157-185-14	Sequence 14, Appl
27	430	20.0	377	3	US-08-281-526B-14	Sequence 14, Appl
28	430	20.0	377	4	US-09-450-790A-14	Sequence 14, Appl
29	430	20.0	377	4	US-09-332-837-14	Sequence 14, Appl
30	430	20.0	377	4	US-09-371-705-2	Sequence 2, Appli
31	430	20.0	422	1	US-07-817-920-3	Sequence 3, Appli
32	430	20.0	422	1	US-08-370-542-3	Sequence 3, Appli
33	430	20.0	422	1	US-08-117-006-3	Sequence 3, Appli
34	430	20.0	422	1	US-08-216-594-3	Sequence 3, Appli
35	430	20.0	422	1	US-08-542-358-3	Sequence 3, Appli
36	430	20.0	422	2	US-08-157-185-13	Sequence 13, Appl
37	430	20.0	422	3	US-08-281-526B-13	Sequence 13, Appl
38	430	20.0	422	3	US-09-018-351-3	Sequence 3, Appli
39	430	20.0	422	4	US-09-332-837-13	Sequence 13, Appl
40	430	20.0	422	5	PCT-US93-00149-3	Sequence 3, Appli
41	429	19.9	422	1	US-07-996-772A-12	Sequence 12, Appl
42	428	19.9	422	4	US-09-341-446B-10	Sequence 10, Appl
43	420.5	19.5	421	4	US-09-450-790A-13	Sequence 13, Appl
44	417	19.4	445	2	US-08-157-185-2	Sequence 2, Appli
45	417	19.4	445	3	US-08-281-526B-2	Sequence 2, Appli

ALIGNMENTS

RESULT 1

US-08-031-538-2
; Sequence 2, Application US/08031538
; Patent No. 5968817
; GENERAL INFORMATION:
; APPLICANT: Sutcliffe, J Gregor
; APPLICANT: Brander, Mark G
; APPLICANT: Lovenberg, Timothy W
; TITLE OF INVENTION: HUMAN SEROTONIN RECEPTORS, DNA ENCODING
; TITLE OF INVENTION: THE RECEPTORS, AND USES THEREOF
; NUMBER OF SEQUENCES: 73
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: The Scripps Research Institute, Office of
; ADDRESSEE: Patent Counsel
; STREET: 10666 No. 5968817th Torrey Pines Road, TPC 8
; CITY: La Jolla
; STATE: CA
; COUNTRY: USA
; ZIP: 92037
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA: US/08/031,538
; APPLICATION NUMBER: 19930315
; FILING DATE: 19930315
; CLASSIFICATION: 800
; ATTORNEY/AGENT INFORMATION:
; NAME: Fitting, Thomas
; REGISTRATION NUMBER: 34,163
; REFERENCE/DOCKET NUMBER: TSP5099P
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 619-554-2937
; TELEFAX: 619-554-6312
; INFORMATION FOR SEQ ID NO: 2:

SEQUENCE CHARACTERISTICS:
 LENGTH: 370 amino acids
 TYPE: AMINO ACID
 TOPOLOGY: linear
 MOLECULE TYPE: protein
 US-08-031-538-2

Alignment Scores:

Pred. No.: 1,248-92 Length: 370
 Score: 1444.50 Matches: 295
 Percent Similarity: 82.8% Conservative: 19
 Best Local Similarity: 77.84% Mismatches: 56
 Query Match: 67.06% Indels: 9
 DB: 2 Gaps: 3

US-09-976-782-15 (1-1152) x US-08-031-538-2 (1-370)

```

QY 5 ATGAGGCGCGTAGCTTTCAGTGGCCACCGCGGCTTCCCTTGGAGCCGAG 64
Db 1 MetGluValCysAsnLeuSerGlyAlaThrProGlyIleAlaPheProProGlyProGlu 20
QY 65 ACCAGAGGAGCCGGGACCCAGACCCGAGAGGATACCTCGTTCGACCCCGAGCGC 124
Db 21 SerCysSer-----AspSerProSerSerGlyArgSerMetGlySerThrProGlyGly 38
QY 125 GCGTCTCTCGCGGCGCGGCGGCGCTTCTCTCTTCCACGTCCTCGTGTGAGCGTG 184
Db 39 LeuIleLeuSerGlyArgGluProProPheSerAlaPheThrValLeuValThrLeu 58
QY 185 CTAGTGTCTGATCGCTGCCACTTTCCTGTGGAACTCTGCTCGGTCCGCTACCATCCG 244
Db 59 LeuValLeuLeuIleAlaAlaThrPheLeuTrpAsnLeuValLeuValThrIleLeu 78
QY 245 CGGTCTCGTCTTCCACCGGTGCGGATACCTTGGTGGCTCGACGCGCTCTCGGAC 304
Db 79 ArgValArgAlaPheHisArgValProHisAsnLeuValAlaSerThrAlaValSerAsp 98
QY 305 GAACATAGTGGCGCTGCGGATGCCACCGAGCTGGCGAGTGAGCTGTCGACCGGCGCA 364
Db 99 ValLeuValAlaAlaLeuValMetProLeuSerLeuValSerGluLeuSerAlaGlyArg 118
QY 365 CGTGGCTCTGCGGCGGAGCTGTGCCACGTGTGATCTCTCTCGACGCGCGAGCTGT 424
Db 119 ArgTrpGlnLeuGlyArgSerLeuCysHisValTrpIleSerPheAspVal----- 135
QY 425 CTGTGTGCGCGCGCGCTCGGAGACGTGGGCGCATCGCTGGCGCGCGGCGGCGC 484
Db 136 LeuCysCysThrAlaSerIleTrpAsnValAlaAlaIleAlaLeuAspArgTrpThr 155
QY 485 ATCACAGGCACCTGCAGCACAGCTGGCCACCGCAGCGCGCTGTGTCTCATGATC 544
Db 156 IleThrArgHisLeuGlnTrpThrLeuArgThrArgArgAlaSerAlaLeuMetIle 175
QY 545 GCGCTCGCGCGGCGTGGCGGTCTATCGCCCTCGCGCGCGCTGTCTTTGCGCGGCGC 604
Db 176 AlaIleThrTrpAlaLeuSerAlaLeuIleAlaLeuAlaProLeuLeuPheGlyTrpGly 195
QY 605 GAGTGTGCGAGCTCGGCTCCAGCGCTGCCAGGTGAGCGCGGAACTCTCTATGCGCGC 664
Db 196 GluAlaTyAspAlaArgLeuGlnArgCysGlnValSerGlnGluProSerTrpAlaVal 215
QY 665 TTCTCCACCGCGCGCTTCCACCGCTGGCGGTGGCTGTGTCTACCGGGAAG 724
Db 216 PheSerThrCysGlyAlaPheTrpValProLeuAlaValValLeuPheValTrpLys 235
QY 725 ATCTACAGCGCGCAAGTTTCGTTTCGCGCGCGCGCGAGAGCTGTGTGCGGTGCGC 784
Db 236 IleTyLeuAlaAlaLysPheArgPheGlyArgArgAlaValValProLeuPro 255
QY 785 GCCACCATGAGTGTGAGTCCAGGTAAAGAGACCTCATGAGCTGAGTGTGTTC 844
Db 256 AlaThrThrGln-----AlaLysGluAlaProGlnGluSerGluThrValPhe 271

```

```

QY 845 ACGGCACATTCGAAAGCAACGGTGTCTTCCAGGTGAGCGGGACTCTCTGGCGGAGCAG 904
Db 272 ThrAlaArgCysArgAlaThrValAlaPheGlnThrSerGlyAspSerTrpArgGln 291
QY 905 AAGGAGGCGGAGCAGCCATGATGTGGGAATTCGATTGGCGTGTGTGTGTGTGTGTGG 964
Db 292 LysGluLysArgAlaAlaMetMetValGlyIleLeuIleGlyValPheValLeuCysTrp 311
QY 965 ATCCCTCTCTCTCTGACGGAACCTCATCAGCCACCTGTGTGCTGCGAGCTGCCCCCATC 1024
Db 312 IleProPhePheLeuThrGluLeuValSerAlaLeuCysAlaCysSerLeuProPhe 331
QY 1025 TGGAAAAGCATATTTCTGTGGCTTGGCTACTCTCAATTTCTTCAACCCCTGATTAC 1084
Db 332 TrpLysSerIlePheLeuTrpLeuGlyTrpSerAsnSerPheAsnProLeuIleTyr 351
QY 1085 ACAGCTTTTAAACAAGACTACAAATGCTTCAGAGCTCTTACTAAGCAGAGA 1141
Db 352 ThrAlaPheAsnLysAsnTrpAsnAsnAlaPheLysSerLeuPheThrLysGlnArg 370

```

RESULT 2

US-08-356-405-2
 ; Sequence 2, Application US/08356405
 ; Patent No. 5807691
 ; GENERAL INFORMATION:
 ; APPLICANT: Amlaiky, No. 5807691rdine
 ; APPLICANT: Boschert, Ursula
 ; APPLICANT: Hen, Rene
 ; APPLICANT: Plassat, Jean-Luc
 ; TITLE OF INVENTION: Polypeptides Having Serotonin Receptor
 ; TITLE OF INVENTION: Activity (5HT5A), Nucleic Acids Coding for These
 ; TITLE OF INVENTION: Polypeptides and Uses Thereof
 ; NUMBER OF SEQUENCES: 13
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Rhone-Poulenc Rorer Inc.
 ; STREET: 500 Arcola Road, 3C43
 ; CITY: Collegeville
 ; STATE: PA
 ; COUNTRY: USA
 ; ZIP: 19426-0107
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: Patent In Release #1.0, Version #1.25
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/356,405
 ; FILING DATE:
 ; CLASSIFICATION: 435
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: WO PCT/FR93/00650
 ; FILING DATE: 29-JUN-1992
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: FR 92/08081
 ; FILING DATE: 01-JUL-1992
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Smith, Julie K
 ; REGISTRATION NUMBER: 38,619
 ; REFERENCE/DOCKET NUMBER: EX92004-US
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: (610)454-3839
 ; TELEFAX: (610)454-3808
 ; INFORMATION FOR SEQ ID NO: 2:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 357 amino acids
 ; TYPE: amino acid
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: protein
 ; US-08-356-405-2

Alignment Scores: 2,17e-63 Length: 357
 Pred. No.: 1022.50 Matches: 219
 Score:


```
QY 500 CAGCACACGCTGCACCGCCGCGCGCTGCTGCTCATGCTCGCGCTCGCGCGGTG 559
Db 148 GltuYrThrLeuArgAlaArgLysArgValSerAsnValMetIleLeuLeuThrTrpAla 167
QY 560 CCGTCGGCTCATGCGCCCTCGCGCGCTGCTCTTTGGCGCGGCGAGGTGTGCAGCGCT 619
Db 168 LeuSerAlaValIleSerLeuAlaProLeuLeuPheGlyTrpGlyGluThrTyrSerGlu 187
QY 620 CGGCTCCAGCGCTGCCAGGTGAGCGCGGACCCCTCTATCGCGCTTCTCCACCGCGGC 679
Db 188 LeuSerGluGluCysGlnValSerArgGluProSerTyrThrValPheSerThrValGly 207
QY 680 GCCTTCCACCTCGCGCTGCGCTGCGCTGCTCTACCGAAGATCTACGAGCGCGCC 739
Db 208 AlaPheTyrLeuProLeuCysValValLeuPheValTyrTrpLysIleTyrLysAlaAla 227
QY 740 AAGTTTCGTTTCGCG---CGCGCGCGAGAGCTGTGCTGCGTGGCGCGCACCATCGAG 796
Db 228 LysPheArgMetGlySerArgLysThrAsnSerValSerProIleProGluAlaValGlu 247
QY 797 GTGAGGTCCAAAGTAAAGGAGACCTGATGAGGCTGAGTGGTGTTCACGGCACATGC 856
Db 248 -----ValLysAspAlaSerGlnHisProGlnMetValPheThrValArg--- 262
QY 857 AAAGCAACGCTGTCTTCCAGGTGAGCGGGGACTCTCGCGGGAGCAGAGGAGCGGA 916
Db 263 HisAlaThrValThrPheGlnThrGluGlyAspThrTrpArgGluGlnLysGluGlnArg 282
QY 917 GCAGCCATGATGTGGGAATCTCATTTGGCGTGTGTGTGCTGTGCTGATCCCTTCTTC 976
Db 283 AlaAlaLeuMetValGlyIleLeuIleGlyValPheValLeuCysTyrPheProPhePhe 302
QY 977 CTGACGGAATCATACGCCCACTGTGCTGCGTGCAGCTGCGCGCGCGCGCGCGCGG 1036
Db 303 ValThrGluLeuIleSerProLeuCysSerTrpAspIleProAlaLeuTrpLysSerIle 322
QY 1037 TTTCTGTGGCTGTGCTACTCAATTTCTTCAACCCCTGATTTACACAGCTTTTAAAC 1096
Db 323 PheLeuTrpLeuGlyTyrSerAsnSerPhePheAsnProLeuIleTyrThrAlaPheAsn 342
QY 1097 AAGAACTACAACTGCTTCAAGAGCGCTCTTTACTAAGCAGAGA 1141
Db 343 ArgSerTyrSerSerAlaPheLysValPhePheSerLysGlnGln 357
```

RESULT 5

```
US-08-461-812-4
; Sequence 4, Application US/08461812
; Patent No. 5935925
; GENERAL INFORMATION:
; APPLICANT: Weinschank, Richard L.
; APPLICANT: Branchek, Theresa
; APPLICANT: Hartig, Paul R.
; TITLE OF INVENTION: DNA ENCODING HUMAN 5-HT1D RECEPTORS AND
; TITLE OF INVENTION: USES THEREOF
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Cooper & Dunham LLP
; STREET: 1185 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: U.S.A.
; ZIP: 10036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/461,812
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: White, John P.
```

```
;
; REGISTRATION NUMBER: 28,678
; REFERENCE/DOCKET NUMBER: 36536-B/JPW/MAT
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 2122780400
; TELEFAX: 2123910525
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 390 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-461-812-4
;
; Alignment Scores:
; Pred. No.: 1,43e-23 Length: 390
; Score: 448.00 Matches: 127
; Percent Similarity: 48.35% Conservative: 63
; Best Local Similarity: 32.32% Mismatches: 147
; Query Match: 20.80% Indels: 56
; DB: 2 Gaps: 13
```

US-09-976-782-15 (1-1152) x US-08-461-812-4 (1-390)

```
QY 59 CCGGAGACACGAGCGGAGCCCGGAGCCCAAGCCCGAGAGGATATCTGGTTTCACCCCG 118
Db 10 ProProProAlaGlySerGluThrTrpValProGlnAlaAsnLeuSerSerAlaPro 29
QY 119 AGC-----GGCGCCGCTCTGCGCGGCGCGG 142
Db 30 SerGlnAsnCysSerAlaLysAspTyrIleTyrGlnAspSerIleSerLeuPro----- 47
QY 143 GGGCGCGCCCTCTCTGCTTTCACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 202
Db 48 -----TrpLysValLeuLeuValMetLeuLeuAlaLeuIleThrLeu 61
QY 203 GCCACTTTCCTGTGGAACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 262
Db 62 AlaThrThrLeuSerAsnAlaPheValIleAlaThrValTyrArgThrArgLysLeuHis 81
QY 263 CGGTCGCGCATAACTTGTGGCTCGACGCGCGCTCTCGGACGAACTAGTGGGAGCGCTG 322
Db 82 ThrProAlaAsnTyrLeuIleAlaSerLeuAlaValThrAspLeuLeuValSerIleLeu 101
QY 323 GCGATGCCACGAGCCTGCGGAGTGAGCTGTGACCGCGGCGAGCTGCGGCTGCTGCGCGG 382
Db 102 ValMetPro-----IleSerThrMetTyrThrValThrGlyArgTrpThrLeuGlyGln 119
QY 383 AGCCTGTGCCACGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 442
Db 120 ValValCysAspPheTrpLeuSerSerAsp-----IleThrCysCysThrAlaSer 136
QY 443 CTCGGGAACTGCGGCGCATGCGCTGCGCGCGGAGCGGCGCATACACGCGACCTGAG 502
Db 137 IleLeuHisLeuCysValIleAlaLeuAspArgTyrTrpAlaIleThrAspAlaValGlu 156
QY 503 CACAGCTGCGCACCGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 562
Db 157 TyrSerAlaLysArgThrProLysArgAlaAlaValMetIleAlaLeuValTrpValPhe 176
QY 563 TCGGCGCTCATGCGCTGCGCGCTGCTGCTGCGCGCGGAGGTGTGCGAGCTCGG 622
Db 177 SerIleSerIleSerLeuProPro---PhePheTrpArgGlnAlaLysAlaGluGluGlu 195
QY 623 CTCAGCGCTGCGAGGTGAGCGCGGAGAACCC---TCCTATCCGCGCTTCTCCACCGCGCG 679
Db 196 ValSerGluCysValValAsnThrAspHisIleLeuTyrThrValTyrSerThrValGly 215
QY 680 GCCTTCCACCTGCGCTTGGCGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 739
Db 216 AlaPheTyrPheProThrLeuLeuIleAlaLeuTyrGlyArgIleTyrValGluAla 235
QY 740 AAGTTTCGTTTC-----GGCGCGCGCG-----CGGAGAGCTGTG 772
```

Db 236 ArgSerArgIleLeuLysGlnThrProAsnArgThrGlyLysArgLeuThrArgAlaGln 255
QY 773 CTG-----CCGTGCGGCGCCACCATCGAGGTCCAAAGCTAAG 814
Db 256 LeuIleThrAspSerProGlySerThrSerSerValThrSerIleAsnSerArgValPro 275
QY 815 GAAGCACCTGATGAGCT---GAAGTGGTGTTCACGGCACATTCGAAAGCAACGGTGTCC 871
Db 276 AspValProSerGluSerGlySerProValTyrValAsnGlnValLysValArgValSer 295
QY 872 TTCAGGTGAGCGGACTCTCGCGGAGCAGAAAG-----GAGAGG 913
Db 296 -----AspAlaLeuLeuGluLysLysLysLeuMetAlaAlaArgGluArg 310
QY 914 CAGCAGCATGATGTTGGGAATTCGATTGGCGTGTGTGCTGCTGCTGATCCCTTC 973
Db 311 LysAlaThrLysThrLeuGlyIleIleLeuGlyAlaPheIleValCysTrpLeuProPhe 330
QY 974 TTCCTGAGGGAACCTCATCAGCCCACTCTGT-----GCCTGCAGGCTGCCCCCATCTGG 1027
Db 331 PheIleIleSerLeuValMetProIleCysLysAspAlaCysTrpPheHisLeuAlaIle 350
QY 1028 AAAAGCATATTTCTGGCTTGGCTTACTCCAAATCTTTCTTCAACCCCTGATTTACACA 1087
Db 351 PheAspPheThrTrpLeuGlyTyrLeuAsnSerLeuIleAsnProIleIleTyrThr 370
QY 1088 GCTTTTAAACAAGAACTACAACATGCTTCAAGAGCCTC 1126
Db 371 MetSerAsnGluAspPheLysGlnAlaPheHisLysLeu 383

RESULT 6

US-08-157-185-15
; Sequence 15, Application US/08157185
; Patent No. 5985585
; GENERAL INFORMATION:
; APPLICANT: Bard A. Jonathan
; APPLICANT: Branchek A. Theresa
; APPLICANT: Weinshank L. Richard
; TITLE OF INVENTION: DNA ENCODING A HUMAN SEROTONIN
; TITLE OF INVENTION: RECEPTOR (5-HT4B) AND USES THEREOF
; NUMBER OF SEQUENCES: 17
; CORRESPONDENCE ADDRESS:
; ADDRESSES: Cooper & Dunham LLP
; CITY: New York
; STATE: New York
; COUNTRY: U.S.A.
; ZIP: 10036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.24
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/157,185
; FILING DATE: 15-JUN-1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: White P., John
; REGISTRATION NUMBER: 28,678
; REFERENCE/DOCKET NUMBER: 41908-A-PCT-US/JPW/MAT
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 278-0400
; TELEFAX: (212) 391-0525
; INFORMATION FOR SEQ ID NO: 15:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 390 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-157-185-15

Alignment Scores:
Pred. No.: 1.43e-23 Length: 390
Score: 448.00 Matches: 127
Percent Similarity: 48.35% Conservative: 63
Best Local Similarity: 32.33% Mismatches: 147
Query Match: 20.80% Indels: 56
DB: 2 Gaps: 13
US-09-976-782-15 (1-1152) x US-08-157-185-15 (1-390)
QY 59 CCCGAGACAGCAGCGGACCCCGGACCCAGAGGAGGATGACTCGTTCGACCCCG 118
Db 10 ProProProAlaGlySerGluThrTrpValProGlnAlaAsnLeuSerSerAlaPro 29
QY 119 AGC-----GGCGCGCTCTCTCGCGGCCGA 142
Db 30 SerGlnAsnCysSerAlaLysAspTyrIleTyrGlnAspSerIleSerLeuPro----- 47
QY 143 GGGCGCCCTTCTCTGTCTTACGGTCTCTGGTGGTGGTGGTCTCTAGTGTCTGTCTG 202
Db 48 -----TriLysValLeuValMetLeuLeuAlaLeuIleThrLeu 61
QY 203 GCCATTTCCTGTGAACCTGCTGCTCGGTTCGATCCATCCCGCGGTCCGTCTCCAC 262
Db 62 AlaThrThrLeuSerAsnAlaPheValIleAlaThrValTyrArgThrArgLysLeuHis 81
QY 263 CGCGTGGCGCATAACTTGTGGCTCGACGCGCGCTCGGACGAACTAGTGGCAGCGCTG 322
Db 82 ThrProAlaAsnTyrLeuIleAlaSerLeuAlaValThrAspLeuValSerIleLeu 101
QY 323 GCGATGCCACCGAGCTGGCGAGTGTGACGCGGCGGACGCTGCTGTGCTGGCGCGG 382
Db 102 ValMetPro-----IleSerThrMetTyrThrValThrGlyArgTrpThrLeuGlyGln 119
QY 383 AGCCTGTGCCAGCTGTGATCTCTTCGACGCGCGGAGCGCTGTCTGTGCTGCCCGCGC 442
Db 120 ValValCysAspPheTrpLeuSerSerAsp-----IleThrCysCysThrAlaSer 136
QY 443 CTCGGGAACGTGGCGCCATCGCCCTGGCGCGCGACGGGCCATCACAGCGCACCTGCAG 502
Db 137 IleLeuHisLeuCysValIleAlaLeuAspArgTyrTrpAlaIleThrAspAlaValGlu 156
QY 503 CACAGCTGCGCACCGCAGCGCGCTGTGTTCATGATCGCGTCCCGCGGTGCGCG 562
Db 157 TyrSerAlaLysArgThrProLysArgAlaAlaValMetIleAlaLeuValTrpValPhe 176
QY 563 TCGGCGCTCATCGCCCTCGCGCGCTGCTTTTGGCGCGGCGGAGGTGTGCGACGCTCGG 622
Db 177 SerIleSerIleSerLeuProPro-----PhePheTrpArgGlnAlaLysAlaGluGlu 195
QY 623 CTCGAGCGTGCAGGTGAGCGCGGAAACCC---TCTATGCGCGCTTCTCCACCGCGCG 679
Db 196 ValSerGluCysValValAsnThrAspHisIleLeuTyrThrValTyrSerThrValGly 215
QY 680 GCCTTCCACCTCCCGTGGCTGGTGTGTTCACCGGAAGATCTACGAGGCGCGC 739
Db 216 AlaPheTyrPheProThrLeuLeuIleAlaLeuTyrGlyArgIleTyrValGluAla 235
QY 740 AAGTTTCGTTTC-----GGCGCGCGC---CGGAGAGGTGTG 772
Db 236 ArgSerArgIleLeuLysGlnThrProAsnArgThrGlyLysArgLeuThrArgAlaGln 255
QY 773 CTG-----CCGTGCGGCGCACCATCGAGGTGAGGTCCAAAGTAAAG 814
Db 256 LeuIleThrAspSerProGlySerThrSerSerValThrSerIleAsnSerArgValPro 275
QY 815 GAAGCACCTGATGAGCT---GAAGTGGTGTTCACGGCACATTCGAAAGCAACGGTGTCC 871
Db 276 AspValProSerGluSerGlySerProValTyrValAsnGlnValLysValArgValSer 295
QY 872 TTCAGGTGAGCGGGAACCTCTCGCGGAGCAGAAAG-----GAGAGG 913
Db 296 -----AspAlaLeuLeuGluLysLysLysLeuMetAlaAlaArgGluArg 310

Qy	563	TCGCGGCTATCGCCCTCGCGCCGCTCTTTGGCCGGGCGGAGTGTCGACGACTCGG	622
Db	177	SerIleSerIleSerLeuProPro---PhePheTrpArgGlnAlaLysAlaGluGluGlu	195
Qy	623	CTCCAGCGCTGCAGGTCGACCGGGAACCC---TCCTATGCGCGCTCTCCACCGCGGC	679
Db	196	ValSerGluCysValValAsnThrAspHisIleLeuThrValTyrSerIleThrValGly	215
Qy	680	GCCTTCACACTGCGCGTGGCGTGGTCCGCTTCTACCGGAAGATCTACGAGCGCGCC	739
Db	216	AlaPheTyrPheProThrLeuLeuLeuAlaLeuTyrGlyArgIleTyrValGluAla	235
Qy	740	AAGTTTCGTTTC-----GGCCGCCGC-----CGGAGAGCTGTG	772
Db	236	ArgSerArgIleLeuLysGlnThrProAsnArgThrGlyLysArgLeuThrArgAlaGln	255
Qy	773	CTG-----CCGTTCCCGGCCACCATGCAGGTGAGGTCCAAGGTAAAG	814
Db	256	LeuIleThrAspSerProGlySerThrSerSerValThrSerIleAsnSerArgValPro	275
Qy	815	GAAGCACTGATGAGGCT---GAAGTGGTGTTCACGGCACATTGCCAAGCAACCGTGTCC	871
Db	276	AspValProSerGluSerGlySerProValTyrValAsnGlnValLysValArgValSer	295
Qy	872	TTCAGGTGAGCGGGACTCCTGCGGGAGCAGAG-----GAGAGG	913
Db	296	-----AspAlaLeuLeuGluLysLysLysLeuMetAlaAlaArgGluArg	310
Qy	914	CGAGCAGCCATGATGGTGGGAATCTGATGGCGGTGTTGTGTGCTGCTCGATCCCTTC	973
Db	311	LysAlaThrLysThrLeuGlyIleLeuLeuGlyAlaPheIleValCysTrpLeuProPhe	330
Qy	974	TTCCTGACGAACTCATCAGCCCACTCTGT-----GCCTGCAGCCTGCCCCCATCTGG	1027
Db	331	PheIleLeuSerLeuValMetProIleCysLysAspAlaCysTrpPheHisLeuAlaIle	350
Qy	1028	AAAAGCATATTTCTGTGGCTTGGCTACTCCAAATTTCTTTCACACCCCTGATTTACACA	1087
Db	351	PheAspPhePheThrTrpLeuGlyTyrLeuAsnSerLeuIleAsnProIleIleTyrThr	370
Qy	1088	GCTTTTAAAGAACTACAACTATGCTTCAAGGCCTC	1126
Db	371	MetSerAsnGluAspPheLysGlnAlaPheHisLysLeu	383
RESULT 9			
US-09-332-837-15			
; Sequence 15, Application US/09332837			
; Patent No. 6432655			
; GENERAL INFORMATION:			
; APPLICANT: Bard, Jonathan			
; APPLICANT: Branchek, Theresa			
; APPLICANT: Weinschank, Richard			
; TITLE OF INVENTION: Methods of Obtaining Pharmaceutical Compositions			
; FILE REFERENCE: 41908-AA-PCI-US			
; CURRENT APPLICATION NUMBER: US/09/332,837			
; CURRENT FILING DATE: 1999-06-14			
; NUMBER OF SEQ ID NOS: 17			
; SOFTWARE: PatentIn version 3.0			
; SEQ ID NO 15			
; LENGTH: 390			
; TYPE: PRT			
; ORGANISM: Homo Sapiens			
US-09-332-837-15			
Alignment Scores:			
Pred. No.:	1,43e-23	Length:	390
Score:	448.00	Matches:	127
Percent Similarity:	48.35%	Conservative:	63
Best Local Similarity:	32.32%	Mismatches:	147
Query Match:	20.80%	Indels:	56
DB:	4	Gaps:	13


```
QY 740 AAGTTTCGTTTC-----GGCGCGCGC---CGGAGAGCTGTG 772
   : : : : :
Db 236 ArgSerArgIleLeuLysGlnThrProAsnArgThrGlyLysArgLeuThrArgAlaGln 255
   : : : : :
QY 773 CTG-----CGTTGCGCGCCACCACGAGTGAGGTGAGGTCCAAAGTAAAG 814
   : : : : :
Db 256 LeuIleThrAspSerProGlySerThrSerSerValThrSerIleAsnSerArgValPro 275
   : : : : :
QY 815 GAAGCACCTGATGAGCT--GAAGTGGTGTTCACGCGCACATTCGAAGCAACGGGTGTC 871
   : : : : :
Db 276 AspValProSerGluSerGlySerProValThrValAsnGlnValLysValArgValSer 295
   : : : : :
QY 872 TTCAGGTGAGCGGGGACTCCTGCGCGGAGCAGAAG-----GAGAGG 913
   : : : : :
Db 296 -----AspAlaLeuLeuGluLysLysLysLeuMetAlaAlaArgGluArg 310
   : : : : :
QY 914 CGAGCAGCCATGATGGTGGGAATTCATGTTGGGTGTTCGTTGCTGTGATCCCTTC 973
   : : : : :
Db 311 LysAlaThrLysThrLeuGlyIleIleLeuGlyAlaPheIleValCysTrpLeuProPhe 330
   : : : : :
QY 974 TTCCTGACGGAACTCATCAGCCACACTGT-----GCCTGACAGCCTGCCCCCATCTCG 1027
   : : : : :
Db 331 PheIleIleSerLeuValMetProIleCysLysAspAlaCysTrpPheHisLeuAlaIle 350
   : : : : :
QY 1028 AAAAGCATATTTCTGTGGCTGGCTACTCCAAATTTCTTCAACCCCTGATTTACACA 1087
   : : : : :
Db 351 PheAspPhePheThrTrpLeuGlyTyrLeuAsnSerLeuIleAsnProIleIleTyrThr 370
   : : : : :
QY 1088 GCTTTTAAACAAGNACTACACAAATGCTTCAAGAGCCTC 1126
   : : : : :
Db 371 MetSerAsnGluAspPheLysGlnAlaPheHisLysLeu 383
```

RESULT 13

```
US-08-216-594-6
; Sequence 6, Application US/08216594
; Patent No. 5652113
; GENERAL INFORMATION:
; APPLICANT: Weinshank, Richard L.
; APPLICANT: Brancheck, Theresa
; APPLICANT: Hartig, Paul R.
; TITLE OF INVENTION: DNA ENCODING A HUMAN 5-HT1F RECEPTOR
; TITLE OF INVENTION: AND USES THEREOF
; NUMBER OF SEQUENCES: 9
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Cooper & Dunham
; STREET: 30 Rockefeller Plaza
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10112
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/216,594
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: White, John P
; REGISTRATION NUMBER: 28,678
; REFERENCE/DOCKET NUMBER: 1795/39318
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-977-9550
; TELEFAX: 212-664-0525
; TELEX: 422523 COOP UI
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 390 amino acids
; TYPE: amino acid
; STRANDEDNESS: unknown
```

```
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; FRAGMENT TYPE: N-terminal
; IMMEDIATE SOURCE:
; CLONE: 5-HT1DB
; US-08-216-594-6
```

Alignment Scores:

```
Pred. No.: 1,56e-22 Length: 390
Score: 433.00 Matches: 124
Percent Similarity: 47.84% Conservative: 64
Best Local Similarity: 31.55% Mismatches: 149
Query Match: 20.10% Indels: 56
DB: 13 Gaps: 13
```

US-09-976-782-15 (1-1152) x US-08-216-594-6 (1-390)

```
QY 59 CCGGAGACCCAGCGAGCCCGGACCCCAAGCCCGAGAGGAGTACTCGTTCGACCCCG 118
   : : : : :
Db 10 ProProAlaProAlaGlySerGluThrTrpValProGlnAlaAsnLeuSerAlaPro 29
   : : : : :
QY 119 AGC-----GGCGCCCTCTCGCGGCCGA 142
   : : : : :
Db 30 SerGlnAsnCysSerAlaLysAspTyrIleTyrGlnAspSerIleSerLeuPro----- 47
   : : : : :
QY 143 GGGCGCCCTCTCTCTCTCTACGGTCTGTGGTGGTGCACGCTGCTAGTGTCTGATCGCT 202
   : : : : :
Db 48 -----TrpLysValLeuLeuValMetLeuLeuAlaLeuIleThrLeu 61
   : : : : :
QY 203 GCACACTTCTCTGGAACCTGCTGCTCGGTCCACCATCCCGGGTCCGTGCTTCCAC 262
   : : : : :
Db 62 AlaThrThrLeuSerAsnAlaPheValIleAlaThrValTyrArgThrArgLysLeuHis 81
   : : : : :
QY 263 CCGGTGCGGCATTAACCTGTGGTTCGACGGCGGTCTCGAGCAAGTACTAGTGGCAGCGCTG 322
   : : : : :
Db 82 ThrProAlaAsnTyrIleIleAlaSerLeuAspValThrAspLeuLeuValSerIleLeu 101
   : : : : :
QY 323 GCGATGCCACCGAGCTGCGGAGTGAGCTGTGCGCGGCGGCGCTGCTGCTGGCGCGG 382
   : : : : :
Db 102 ValIlePro-----IleSerThrMetTyrThrValThrAspArgTrpThrLeuSerGln 119
   : : : : :
QY 383 AGCTGTGTCACGTGTGATCTCTTCGACGCGCGGAGCCTGTCTGTGCTGCCCGCGCGC 442
   : : : : :
Db 120 ValValCysAspPheTrpLeuSerSerAsp-----IleThrCysCysThrAlaSer 136
   : : : : :
QY 443 CTCGGGAAGTGGCGCCATCGCCCTGGCGCGGAGCGGCGCATCACAGGACCTGCAG 502
   : : : : :
Db 137 IleLeuHisLeuCysValIleAlaLeuAspArgTyrTrpAlaIleThrAspAlaValGlu 156
   : : : : :
QY 503 CACAGCTGCGCACCGCGCGCTGTTGCTCATGATCGCTGCTGCTGCTGCTGCTGCTGCTG 562
   : : : : :
Db 157 TyrSerAlaLysArgThrProLysArgAlaAlaValMetIleAlaLeuValTrpValPhe 176
   : : : : :
QY 563 TCGCGCTCATCGCCCTCGCGCGCTGCTGCTTTTGGCGGGCGGAGGTGTGCGACGCTCG 622
   : : : : :
Db 177 SerIleSerIleSerLeuProPro---PhePheTrpArgGlnAlaLysAlaGluGlu 195
   : : : : :
QY 623 CTCGAGCGCTGCGAGTGAGCGGGAACCC---TCCTATGCGCCTCTTCCACCGCGCG 679
   : : : : :
Db 196 ValSerGluCysValValAsnThrAspHisIleLeuTyrThrValTyrSerThrValGly 215
   : : : : :
QY 680 GCTTTCACCTCGCGCTTGGCGGTGTGCTTGTCTACCGAAGATCTACGAGCGCGC 739
   : : : : :
Db 216 AlaPheTyrPheProThrLeuLeuIleAlaLeuTyrGlyArgIleTyrValGluAla 235
   : : : : :
QY 740 AAGTTTCGTTTC-----GGCGCGCGC---CGGAGAGCTGTG 772
   : : : : :
Db 236 ArgSerArgIleLeuLysGlnThrProAsnArgThrGlyLysArgLeuThrArgAlaGln 255
   : : : : :
QY 773 CTG-----CCGTTGCGCGCACCATCATGAGTGTAGGTCGAGGTAAAG 814
   : : : : :
```


Db 296 -----AspAlaLeuLeuGluLysLysLeuMetAlaAlaArgGluArg 310
Qy 914 CGAGCAGCATCATGGTGGGAATTCGATTGGCGCTGTTGTGCTGCTGGATCCCTTC 973
Db 311 LysAlaThrLysThrLeuGlyIleLeuGlyAlaPheIleValCysTrpLeuProPhe 330
Qy 974 TTCCTGAGGAACTCATCAGCCCACTCTGT-----GCCTGCAGCGCTGCCCCCACTCTGG 1027
Db 331 PheIleIleSerLeuValMetProIleCysLysAspAlaCysTrpPheHisLeuAlaIle 350
Qy 1028 AAAAGCATATTTCTGTGGCTTGGCTACTCAATCTTTCTTCAACCCCTGATTACACA 1087
Db 351 PheAspPheThrTrpLeuGlyTyrLeuAsnSerLeuIleAsnProIleIleTyrThr 370
Qy 1088 GCTTTTACAGAACTACACAACTGCCTCAAGAGCCTC 1126
Db 371 MetSerAsnGluAspPheLysGlnAlaPheHisLysLeu 383

RESULT 15

US-08-370-542-6
Sequence 6, Application US/08370542
Patent No. 5476782
GENERAL INFORMATION:
APPLICANT: Weinshank, Richard L.
APPLICANT: Brancheck, Theresa
APPLICANT: Hartig, Paul R.
TITLE OF INVENTION: DNA ENCODING A HUMAN 5-HT1E RECEPTOR AND
TITLE OF INVENTION: USES THEREOF
NUMBER OF SEQUENCES: 12
CORRESPONDENCE ADDRESS:
ADDRESSER: Cooper & Dunham
STREET: 30 Rockefeller Plaza
CITY: New York
STATE: New York
COUNTRY: U.S.A.
ZIP: 10112
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/370,542
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/194,113
FILING DATE:
APPLICATION NUMBER: US/07/803,626
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: White, John P.
REGISTRATION NUMBER: 28,678
REFERENCE/DOCKET NUMBER: 1795/39317
TELECOMMUNICATION INFORMATION:
TELEPHONE: 212-977-9550
TELEFAX: 212-664-0525
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 398 amino acids
TYPE: amino acid
STRANDEDNESS: unknown
TOPOLOGY: unknown
MOLECULE TYPE: protein
HYPOTHETICAL: NO
ANTI-SENSE: NO
FRAGMENT TYPE: N-terminal
US-08-370-542-6

Alignment Scores: 1.57e-22 Length: 398
Pred. No.: 433.00 Matches: 124
Score:

Percent Similarity: 47.84% Conservative: 64
Best Local Similarity: 31.55% Mismatches: 149
Query Match: 20.10% Indels: 56
DB: 1 Gaps: 13
US-09-976-782-15 (1-1152) x US-08-370-542-6 (1-398)
Qy 59 CCGAGAGACCAGCAGCGGACCCCGGACCCCAAGCCCGAGAGGATCTCGTTTCGACCCCG 118
Db 10 ProProAlaProAlaGlySerGluThrTrpValProGlnAlaAsnLeuSerSerAlaPro 29
Qy 119 AGC-----GGCGCCGCTCTCGCGGCCGA 142
Db 30 SerGlnAsnCysSerAlaLysAspTyrIleTyrGlnAspSerIleSerLeuPro----- 47
Qy 143 GGGCGCCCTCTCTGTCTTACGGCTGCTGGTGGTGACGCTGCTAGTGTGCTGATCGCT 202
Db 48 -----TrpLysValLeuLeuValMetLeuLeuAlaLeuIleThrLeu 61
Qy 203 GCCACTTCTCTGTGAACCTGCTGCTCCGCTCACCATCCCGCGGCTCGCTTCCAC 262
Db 62 AlaThrThrLeuSerAsnAlaPheValIleAlaThrValTyrArgThrArgLysLeuHis 81
Qy 263 CCGGTGCGCATAACTTGTGGCTCGACGGCGCTCTCGGACGAACTAGTGGCAGCGCTG 322
Db 82 ThrProAlaAsnTyrLeuIleAlaSerLeuAspValThrAspLeuLeuValSerIleLeu 101
Qy 323 CGGATGCCACCGAGCTGGCGAGTGTGTGCGCGGCGGCTGCGCTGGGCGCG 382
Db 102 ValIlePro-----IleSerThrMetTyrThrValThrAspArgTrpThrLeuSerGln 119
Qy 383 AGCCTGTGTCACGTGTGATCTTCTTCAGACCGCGGAGCTCTGTGTGCTGCCCGCGCG 442
Db 120 ValValCysAspPheTrpLeuSerSerAsp-----IleThrCysCysThrAlaSer 136
Qy 443 CTCGGGAACGTGGCGGCATCGCCCTGGCGCGGACGGGGCCATCACAGCGCACCTGCGAG 502
Db 137 IleLeuHisLeuCysValIleAlaLeuAspArgTyrTrpAlaIleThrAspAlaValGlu 156
Qy 503 CACAGCTGCGCACCGCGCGCGCTGTTGTCTCATGATCGCGCTCGCCCGGGTGGCG 562
Db 157 TyrSerAlaLysArgThrProLysArgAlaAlaValMetIleAlaLeuValTrpValPhe 176
Qy 563 TCGGGGCTCATCGCCCTGCGCGCTGCTTTTGGCGGGGAGGTGTGCGACGCTCG 622
Db 177 SerIleSerIleSerLeuProPro---PhePheTrpArgGlnAlaLysAlaGluGlu 195
Qy 623 CTCAGCGCTGCCAGGTGAGCGGGAACCC---TCGTATGCGCCTTCTCCACCCCGCG 679
Db 196 ValSerGluCysValValAsnThrAspHisIleLeuTyrThrValTyrSerThrValGly 215
Qy 680 GCCTTCCACCTGCGCTGGCTGGTGGCTTGTCTACCGGAAGATCTACGAGGCGCG 739
Db 216 AlaPheTyrPheProThrLeuLeuIleAlaLeuTyrGlyArgIleTyrValGluAla 235
Qy 740 AAGTTTCGTTTC-----GGCGCGCGC---CGGAGAGCTGTG 772
Db 236 ArgSerArgIleLeuLysGlnThrProAsnArgThrGlyLysArgLeuThrArgAlaGln 255
Qy 773 CTG-----CCGTGCGCGGCACCATCGAGTGAGGTCCAGGTAAAG 814
Db 256 LeuIleThrAspSerProGlySerThrSerSerValThrSerIleAsnSerArgValPro 275
Qy 815 GAAGCACCTGATGAGGCT---GAAGTGGTGTTCAGGCACTATTGCAAAACACGCTGCTCC 871
Db 276 AspValProSerGluSerGlySerProValTyrValAsnGlnValLysValArgValSer 295
Qy 872 TTCCAGGTGAGCGGGACTCTCGCGGGGAGCAGAG-----GAGAGG 913
Db 296 -----AspAlaLeuLeuGluLysLysLeuMetAlaAlaArgGluArg 310
Qy 914 CGAGCAGCCATGATGTTGGGAATTCGATTGGCGTGTGTTGTGCTGCTGATCCCTTC 973
Db 124

Db 311 LysAlaThrLysThrLeuGlyIleIleLeuGlyAlaPheIleValCysTrpLeuProPhe 330
Qy 974 TTCTGTACGGAACTCATCAGCCCACTCTGT-----GCCTGCAGCCTGCCCCCACTCTGG 1027
Db 331 PheIleIleSerLeuValMetProIleCysLysAspAlaCysTrpPheHisLeuAlaIle 350
Qy 1028 AAAAGCATATTCTGTGGCTTGCTACTCCCAATTCCTTCAACCCCTGATTACACA 1087
Db 351 PheAspPheThrTrpLeuGlyTyLeuAsnSerLeuIleAsnProIleIleTyThr 370
Qy 1088 GCTTTTAACAAGAACTACAACAATGCCTTCAAGAGCCTC 1126
Db 371 MetSerAsnGluAspPheLysGlnAlaPheHisLysLeu 383

Search completed: August 31, 2004, 20:24:14
Job time : 47 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - protein search, using frame_plus_n2p model

Run on: August 31, 2004, 19:59:22 ; Search time 28.5 seconds

(without alignments)
4209.461 Million cell updates/sec

Title: US-09-976-782-15

Perfect score: 2154

Sequence: 1 cgcagtgaggccgtagcc.....aagcagagatgaacacaggg 1152

Scoring table:

BLOSUM62
Xgapop 10.0, Xgapext 0.5
Ygapop 10.0, Ygapext 0.5
Fgapop 6.0, Fgapext 7.0
Delop 6.0, Delext 7.0

Searched: 141681 seqs, 52070155 residues

Total number of hits satisfying chosen parameters: 283362

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+n2p.model p/US09976782/runat 31082004 115950 10303/app_query.fasta 1.1351
-Q=/cgn2 1/USPTO spool p/US09976782/runat 31082004 115950 10303/app_query.fasta 1.1351
-DB=SwissProt 42 -OFMT=fastan -SUFFIX=n2p.rsp -MINMATCH=0.1 -LOOPCL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi
-LIST=45 -DOALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=pct -NORM=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09976782 @CGN 1 1 35 @runat 31082004 115950 10303 -NCPU=6 -ICPU=3
-NO MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : SwissProt 42:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query %	Match	Length	DB ID	Description
1	1453.5	67.5	370	1	SH5B RAT	P35365 rattus norv
2	1452.5	67.4	370	1	SH5B_MOUSE	P31387 mus musculu
3	1063.5	49.4	357	1	SH5A HUMAN	P47898 homo sapien
4	1022.5	47.5	357	1	SH5A_MOUSE	P30966 mus musculu
5	1014.5	47.1	357	1	SH5A RAT	P35364 rattus norv
6	448	20.8	390	1	SH1B HUMAN	P28222 homo sapien
7	448	20.8	390	1	SH1B_PANTR	P60020 pan troglod
8	447	20.8	386	1	SH1B_SPAEH	P56496 spalax leuc
9	444.5	20.6	389	1	SH1B_CAVPO	O08892 cavia porce
10	444	20.6	390	1	SH1B_RABIT	P49144 oryctolagus
11	441.5	20.5	564	1	SH1I DROME	P20905 drosophila
12	441	20.5	377	1	SH1D CANFA	P11614 canis fami
13	440	20.4	386	1	SH1B_CRIGR	P46636 cricetulus
14	438	20.3	386	1	SH1B_MOUSE	P28334 mus musculu
15	437.5	20.3	376	1	SH1D_CAVPO	Q60484 cavia porce
16	437	20.3	386	1	SH1B RAT	P28564 rattus norv
17	435.5	20.2	416	1	SH1B_FUGRU	O42384 fugu rubrip
18	430	20.0	377	1	SH1D_HUMAN	P28221 homo sapien

19	429.5	19.9	446	1	SHT_BOMMO	Q17239 bombyx mori
20	428	19.9	422	1	SH1A_HUMAN	P08908 homo sapien
21	428	19.9	422	1	SH1A_PANTR	P08908 pan troglod
22	425.5	19.8	379	1	SH1D_FUGRU	P79748 fugu rubrip
23	425	19.7	421	1	SH1A_MOUSE	Q64264 mus musculu
24	421	19.5	422	1	SH1A RAT	P19327 rattus norv
25	419.5	19.5	374	1	SH1D_MOUSE	Q61224 mus musculu
26	418	19.4	466	1	SHT_HELVI	Q25190 heliothis v
27	417	19.4	479	1	SH7_HUMAN	Q25414 lymnaea sta
28	416.5	19.3	509	1	SHT_LYMST	P34969 homo sapien
29	416	19.3	374	1	SH1D RAT	P28565 rattus norv
30	416	19.3	377	1	SH1D_RABIT	P49145 oryctolagus
31	416	19.3	446	1	SH7_CAVPO	P50407 cavia porce
32	413.5	19.2	366	1	SH1F_CAVPO	O08890 cavia porce
33	412.5	19.2	366	1	SH1F_HUMAN	P30939 homo sapien
34	412.5	19.2	388	1	SH1B_DIDMA	P35404 didelphis m
35	411.5	19.1	365	1	SH1F_PANTR	Q952d9 pan troglod
36	411.5	19.1	366	1	SH1F_MOUSE	Q02284 mus musculu
37	409.5	19.0	379	1	GRE2_BALAM	Q93127 balanus amp
38	408	18.9	363	1	SH1E_PANTR	Q9n2b6 pan troglod
39	408	18.9	365	1	SH1E_HUMAN	P28566 homo sapien
40	408	18.9	448	1	SH7 RAT	P32305 rattus norv
41	408	18.9	484	1	OAR1_LOCMI	Q25321 locusta mig
42	408	18.9	484	1	OAR2_LOCMI	Q25322 locusta mig
43	406.5	18.9	366	1	SH1F RAT	P30940 rattus norv
44	406.5	18.9	423	1	SH1A_FUGRU	O42385 fugu rubrip
45	403	18.7	448	1	SH7_MOUSE	P32304 mus musculu

ALIGNMENTS

RESULT 1
SH5B RAT STANDARD; PRT; 370 AA.
ID P35365;
AC 01-JUN-1994 (Rel. 29, Created)
DT 01-JUN-1994 (Rel. 29, Last sequence update)
DT 01-OCT-1996 (Rel. 34, Last annotation update)
DE 5-hydroxytryptamine 5B receptor (5-HT-5B) (Serotonin receptor) (MR22).
GN HTR5B OR 5HT5B.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Sprague-Dawley; TISSUE=Brain;
RX MEDLINE=93234515; PubMed=7682702;
RA Erlander M.G., Lovenberg T.W., Baron B.M., de Lecea L.,
RA Danielson P.E., Racke M., Slone A.L., Siegel B.W., Foye P.E.,
RA Cannon K., Burns J.E., Sutcliffe G.J.;
RT "Two members of a distinct subfamily of 5-hydroxytryptamine receptors
differentially expressed in rat brain.";
RL Proc. Natl. Acad. Sci. U.S.A. 90:3452-3456(1993).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=94039744; PubMed=8224165;
RA Wisden W., Parker E.M., Mahle C.D., Grisel D.A., Nowak H.P.,
RA Yocca F.D., Felder C.C., Seeburg P.H., Voigt M.M.;
RT "Cloning and characterization of the rat 5-HT5B receptor. Evidence
that the 5-HT5B receptor couples to a G protein in mammalian cell
membranes.";
RL FEBS Lett. 333:25-31(1993).
CC -!- FUNCTION: This is one of the several different receptors for 5-
hydroxytryptamine (serotonin), a biogenic hormone that functions
as a neurotransmitter, a hormone, and a mitogen. The activity of
this receptor is mediated by G proteins. Probably involved in
anxiety and depression.
CC -!- SUBCELLULAR LOCATION: Integral membrane protein.
CC -!- TISSUE SPECIFICITY: Brain; in the CA1 region of hippocampus, the
medial habenula, and raphe nuclei.
CC -!- SIMILARITY: Belongs to family 1 of G-protein coupled receptors.
CC -!- STRONGEST TO THE OTHER 5HT-5 SUBTYPE RECEPTORS.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----

DR ENBL; L10073; AAA40616.1; --
 DR PIR; S38744; S38744;
 DR InterPro; IPR00276; GPCR_Rhodopsin.
 DR Pfam; PF00001; 7tm_1; 1.
 DR PRINTS; PR00237; GPCR_Rhodopsin.
 DR PROSITE; PS00237; G-PROTEIN RECEPTOR FL 1; 1.
 DR PROSITE; PS00262; G-PROTEIN RECEPTOR FL 2; 1.
 KW G-protein coupled receptor; Transmembrane; Glycoprotein;
 KW Multigene family.
 FT DOMAIN 1 52
 FT TRANSMEM 53 75
 FT DOMAIN 76 90
 FT TRANSMEM 91 111
 FT DOMAIN 112 128
 FT TRANSMEM 129 150
 FT DOMAIN 151 171
 FT TRANSMEM 172 194
 FT DOMAIN 195 211
 FT TRANSMEM 212 232
 FT DOMAIN 233 295
 FT TRANSMEM 296 316
 FT DOMAIN 317 333
 FT TRANSMEM 334 354
 FT DOMAIN 355 370
 FT CARBOHYD 5 5
 FT DISULFID 127 205
 SQ SEQUENCE 370 AA; 41122 MW; 8ECSF79BFD647E5 CRC64;

Alignment Scores:

Pred. No.: 3,256-67 Length: 370
 Score: 1453.50 Matches: 296
 Percent Similarity: 83.38% Conservative: 20
 Best Local Similarity: 78.10% Mismatches: 54
 Query Match: 67.48% Indels: 9
 DB: 1 Gaps: 3

US-09-976-782-15 (1-1152) x 5HSB_RAT (1-370)

QY 5 ATGGAGCGCTAGCTTTCAGTGGCCACCGCGCGGTTCCTTCCCTGGGACCCGAG 64
 Db 1 MetGluValSerAsnLeuSerGlyAlaThrProGlyLeuAlaPheProGlyProGlu 20
 QY 65 ACCAGCAGCGGACCCCGGACCCCAAGCCGAGAGGATACCTCGGTTTCGACCCCGAGCGC 124
 Db 21 SerCysSer-----AspSerProSerSerGlyArgSerMetGlySerThrProGlyGly 38
 QY 125 GCGTCTCGCGGCGCGAGGCGCGCTTCTCTGTCTTTCACGCTCTCGTGTGTGAGCTG 184
 Db 39 LeuValLeuSerGlyArgGluProPheSerAlaPheThrValLeuValThrLeu 58
 QY 185 CTAGTCTGCTGATCGCTGCGCACTTCTCTGTGGAACCTGCTGGTTCGGTTCACATCCCG 244
 Db 59 LeuValLeuLeuLeuAlaAlaThrPheLeuTrpAsnLeuValValThrLeu 78
 QY 245 CGGTCTCGCTTCCACCGCGTTCGCGCATACCTCGTGGCTTCGACGCGCTCTCGGAC 304
 Db 79 ArgValArgAlaPheHisArgValProHisAsnLeuValAlaSerThrAlaValSerAsp 98
 QY 305 GAACATAGTGAGCGCTGCGGATGCCACCGAGCTCGCGAGTGTGACCGCGGCA 364
 Db 99 ValLeuValAlaAlaLeuValMetProLeuSerLeuValSerGluLeuSerAlaGlyArg 118
 QY 365 CQTGCGTCTGCGGCGGAGCTGTGCGACGTGTGATCTCTCTCGACCGCGAGCTGT 424

Db 119 ArgTrpGlnLeuGlyArgSerLeuCysHisValTrpIleSerPheAspVal----- 135
 QY 425 CTGTGCTGCGCGCGCGCTCGGGAACGTGGCGGCATTCGCTGGGCGCGACGGGCGC 484
 Db 136 LeuCysCysThrAlaSerIleTrpAsnValAlaAlaIleAlaLeuAspArgTyrTrpThr 155
 QY 485 ATCACAGGACCTGACGACACGCTCGCGACCGCGAGCGCGCTCGTTCCTCATGTC 544
 Db 156 IleThrArgHisLeuGlnTyrThrLeuArgThrArgArgAlaSerAlaLeuMetIle 175
 QY 545 GCGCTCGCGCGCGGTGCGCTCATCGCCCTCGCGCGCTGCTCTTTGGCGGGCGC 604
 Db 176 AlaIleThrTrpAlaLeuSerAlaLeuLeuAlaLeuAlaProLeuLeuPheGlyTrpGly 195
 QY 605 GAGTGTGCGACGCTCGCTCGCGCTCGCGAGTGCAGCGGGAACCTCTCTATGCGCGC 664
 Db 196 GluAlaTyrAspAlaArgLeuGlnArgCysGlnValSerGlnGluProSerTyrAlaVal 215
 QY 665 TTCTCCACCGCGCGCGCTTCCACCTCGCGCTTGGCGTGGTGGCTGTGTCTACCGGAG 724
 Db 216 PheSerThrCysGlyAlaPheTyrValProLeuAlaValValLeuPheValTyrTrpPhe 235
 QY 725 ATCTACGAGCGCGCAAGTTTCGTTTCGCGCGCGCGAGAGCTGTGCTGCGCTTCGCG 784
 Db 236 IleTyrLeuAlaAlaLysPheArgPheGlyArgArgArgAlaValValProLeuPro 255
 QY 785 GCCACCATGAGGTGAGTCCCAAGGTAAAGAGACCTGATGAGGCTGAAGTGGTTC 844
 Db 256 AlaThrThrGln-----AlaLysGluAlaProGlnGluSerGluThrValPhe 271
 QY 845 ACCGCACTTGCACCAAGCGGTCTCTCCAGTGCAGCGGAGCTGCTCGCGGAGCAG 904
 Db 272 ThrAlaArgCysArgAlaThrValAlaPheGlnThrSerGlyAspSerTrpArgGluGln 291
 QY 905 AAGCAGAGCGCAGCAGCCATGATGTGGGAATTCATGTCGCTGTGTGTGTGTGTG 964
 Db 292 LysGluLysArgAlaAlaMetMetValGlyIleLeuIleGlyValPheValLeuCysTrp 311
 QY 965 ATCCCTCTTCTCTGACGAACTCATCAGCCACTCTGTGCTGCGCTGCGCCCGCCATC 1024
 Db 312 IleProPhePheLeuThrGluLeuValSerProLeuCysAlaCysSerLeuProIle 331
 QY 1025 TGGAAAAGCATATTCTGTGGCTTGGCTACTCTCAATTCCTTCTTCAACCCCGATTAC 1084
 Db 332 TrpLysSerIlePheLeuTrpLeuGlyTyrSerAsnSerPhePheAsnProLeuIleTyr 351
 QY 1085 ACAGCTTTTAAAGAACTACAAATGCTTCAAGAGCTCTTTTACTAAGCAGAGA 1141
 Db 352 ThrAlaPheAsnLysAsnTyrAsnAsnAlaPheLysSerLeuPheThrLysGlnArg 370

RESULT 2

5HSB_MOUSE STANDARD; PRT; 370 AA.
 ID 5HSB_MOUSE
 AC P31387;
 DT 01-JUL-1993 (Rel. 26, Created)
 DT 01-JUL-1993 (Rel. 26, Last sequence update)
 DT 15-MAR-2004 (Rel. 43, Last annotation update)
 DE 5-hydroxytryptamine 5B receptor (5-HT-5B) (Serotonin receptor).
 GN HTR5B OR 5HT5B.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]_TaxID=10090;
 RP SEQUENCE FROM N.A.
 RC TISSUE=Brain;
 RX MEDLINE=93196607; PubMed=8450829;
 RA Matthes H., Boschert U., Amlaiky N., Grailhe R., Plassat J.-L.,
 RA Muscatelli F., Mattei M.-G., Hen R.;
 RT "Mouse 5-hydroxytryptamine5A and 5-hydroxytryptamine5B receptors
 RT define a new family of serotonin receptors: cloning, functional
 RT expression, and chromosomal localization.";
 RL Mol. Pharmacol. 43:313-319(1993).

CC -i- FUNCTION: This is one of the several different receptors for
CC 5-hydroxytryptamine (serotonin), a biogenic hormone that functions
CC as a neurotransmitter, a hormone, and a mitogen. The activity of
CC this receptor is mediated by G proteins. Probably involved in
CC anxiety and depression.
CC -i- SUBCELLULAR LOCATION: Integral membrane protein.
CC -i- TISSUE SPECIFICITY: Expressed predominantly in the central nervous
CC system; in the hippocampus, habenula, and the dorsal raphe.
CC -i- SIMILARITY: Belongs to family 1 of G-protein coupled receptors.
CC STRONGEST TO THE OTHER 5HT-5 SUBTYPE RECEPTORS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: X69867; CAA49501.1; --
CC PIR: I48231; I48231.
CC MGD: MGI:96284; Htr5b.
CC DR InterPro: IPR000276; GPCR_Rhodopsn.
CC DR Pfam: PF00001; 7tm 1; 1.
CC DR PRINTS: PR00237; GPCRHHODOPSN.
CC DR PROSITE: PS00237; G-PROTEIN RECEPTOR F1.1; 1.
CC DR PROSITE: PS00262; G-PROTEIN RECEPTOR F1.2; 1.
CC KW G-protein coupled receptor; Transmembrane; Glycoprotein;
CC Multigene family.
CC FT DOMAIN 1 52 EXTRACELLULAR (POTENTIAL).
CC FT TRANSMEM 53 75 1 (POTENTIAL).
CC FT DOMAIN 76 90 CYTOPLASMIC (POTENTIAL).
CC FT TRANSMEM 91 111 2 (POTENTIAL).
CC FT DOMAIN 112 128 EXTRACELLULAR (POTENTIAL).
CC FT TRANSMEM 129 150 3 (POTENTIAL).
CC FT DOMAIN 151 171 CYTOPLASMIC (POTENTIAL).
CC FT TRANSMEM 172 194 4 (POTENTIAL).
CC FT DOMAIN 195 211 EXTRACELLULAR (POTENTIAL).
CC FT TRANSMEM 212 232 5 (POTENTIAL).
CC FT DOMAIN 233 295 CYTOPLASMIC (POTENTIAL).
CC FT TRANSMEM 296 316 6 (POTENTIAL).
CC FT DOMAIN 317 333 EXTRACELLULAR (POTENTIAL).
CC FT TRANSMEM 334 354 7 (POTENTIAL).
CC FT DOMAIN 355 370 CYTOPLASMIC (POTENTIAL).
CC FT CARBOHYD 5 5 N-LINKED (GLCNAC. . .) (POTENTIAL).
CC FT DISULFID 127 205 BY SIMILARITY.
CC SQ SEQUENCE 370 AA; 41201 MW; 0553C62B12DAAD84 CRC64;

Alignment Scores:
Pred. No.: 3,66e-67 Length: 370
Score: 1452.50 Matches: 298
Percent Similarity: 83.84% Conservative: 19
Best Local Similarity: 78.63% Mismatches: 53
Query Match: 67.43% Indels: 9
DB: 1 Gaps: 3

US-09-976-782-15 (1-1152) x 5H5B_MOUSE (1-370)

Qy 5 ATGGAGGCGCGTAGCTTTCAGTGGCCACCGCGCGGTGCTGCTGCGCTGGGACCGCGAG 64
Db 1 MetGluValSerAsnLeuSerGlyAlaThrProGlyLeuAlaPheProProGlyProGlu 20
Qy 65 ACCGACGCGGACCGGACCGGACCGGAGGAGGACTCGGTTCCAGCCCGGCGGC 124
Db 21 SerCysSer-----AspSerProSerSerGlyArgSerMetGlySerThrProGly 38
Qy 125 GCGGTCCTGCGGCGGAGGCGCGCTTCTGCTTCAGGTCCTCGTGGTGGACGTG 184
Db 39 LeuIleLeuProGlyArgGluProProPheSerAlaPheThrValValValThrLeu 58
Qy 185 CTAGTGTGCTGATCGCTGCCACTTCTCTGTGGAACTGCTGGTCCGCTCACCATCCG 244
Db 59 LeuValLeuLeuAlaAlaThrPheLeuTrpAsnLeuValLeuValThrIleLeu 78

Qy 245 CGGTCGCTGCTTCCACCGCGCTGCCGATACCTGCTGGCTCGACGCGCGCTCTCGGAC 304
Db 79 ArgValAlaAlaPheHisArgValProHisAsnLeuValAlaSerThrAlaValSerAsp 98
Qy 305 GAACCTAGTGGCAGCGCTGGCGATGCCACCGAGCTGGCGAGTGAGTGTTCGACCGGCGA 364
Db 99 ValLeuValAlaValLeuValMetProLeuSerLeuValSerGluLeuSerAlaGlyArg 118
Qy 365 CGTCGCTGCTGGCGCGGAGCTGTGCCAGTGTGGATCTCTTCGACGCGGAGCGCTGT 424
Db 119 ArgTrpGlnLeuGlyArgSerLeuCysHisValTrpIleSerPheAspVal----- 135
Qy 425 CTGTGCTGCCCGCGCGCTCGGGAACGTGGCGCATCGCCCTCGCGCGCGCGCGCGGC 484
Db 136 LeuCysCysThrAlaSerIleTrpAsnValAlaAlaIleAlaLeuAspArgTyrTrpThr 155
Qy 485 ATCACGCGCACCTGCAGCACAGCTCGCACCGCGAGCGCGCTCTGTGCTCATGATC 544
Db 156 IleThrArgHisLeuGlnTyrThrLeuArgThrArgSerArgAlaSerAlaLeuMetIle 175
Qy 545 GCGCTGCCCGGCTGGCGCTCGCGCTCATCGCCCTCGCGCGCTCTTGGCGGCGGC 604
Db 176 AlaIleThrTrpAlaLeuSerAlaLeuIleAlaLeuAlaProLeuLeuPheGlyTrpGly 195
Qy 605 GAGGTGGCGAGCGCTCGGCTCCAGCGCTGCCAGGTGAGCGGGGAACCTCTCTATGCCGCC 664
Db 196 GluAlaTyrAspAlaArgLeuGlnArgCysGlnValSerGlnGluProSerTyrAlaVal 215
Qy 665 TTCTCCACCGCGCGCTTCCACCTCGCGCTGGCGCTGGCGCTGTGCTTGTACCGGAAG 724
Db 216 PheSerThrCysGlyAlaPheTyrLeuProLeuAlaValValLeuPheValTyrTrpLys 235
Qy 725 ATCTACGAGCGGCGCAAGTTTCGTTTCGCGCGCGCGCGAGAGCTGTGCTGCCCTCCG 784
Db 236 IleTyrLysAlaAlaLysPheArgPheGlyArgArgAlaValAlaValProLeuPro 255
Qy 785 GCCACCATGAGGTGAGTCCAGGTAAAGAACACCTGATGAGGTGAGTGGTGTTC 844
Db 256 AlaThrThrGln-----AlaLysGluAlaProProGluSerGluMetValPhe 271
Qy 845 ACGGCACATTGCAAGCAACAGCTGCTCTCCAGGTGAGCGGGGACTCTCTGCGGAGAGCAG 904
Db 272 ThrAlaArgArgAlaThrValThrPheGlnThrSerGlyAspSerTrpArgGluGln 291
Qy 905 AAGGAGGCGGAGCAGCATGCTGGGAATCTGATGCGCTGTGCTGTGCTGCTGG 964
Db 292 LysGluLysArgAlaAlaMetMetValGlyIleLeuIleGlyValPheValLeuCysTrp 311
Qy 965 ATCCCTCTTCTCTGACGGAACCTCATCAGCCCACTCTGTGCTGCGAGCTGCCCGCCATC 1024
Db 312 IleProPhePheLeuThrGluLeuIleSerProLeuCysAlaCysSerLeuProPhe 331
Qy 1025 TGGAAAAGCATATTTCTGTGCTGGCTACTCCAAATTTCTTCAACCCCTGATTAC 1084
Db 332 TrpLysSerIlePheLeuTrpLeuGlyTyrSerAsnSerPhePheAsnProLeuIleTyr 351
Qy 1085 ACAGCTTTTAAACAGAACTACAATGCTCTTCAAGACCTCTTACTAAGCAGAGA 1141
Db 352 ThrAlaPheAsnLysAsnTyrAsnAsnAlaPheLysSerLeuPheThrLysGlnArg 370

RESULT 3

ID SH5A_HUMAN STANDARD; PRT; 357 AA.
AC P47898;
DT 01-FEB-1996 (Rel. 33, Created)
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE 5-hydroxytryptamine 5A receptor (5-HT-5A) (Serotonin receptor)
GN (5-HT-5).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=95080386; PubMed=7998681;
 RA Rees S., den Baas I., Foord S., Goodson S., Bull D., Kilpatrick G.,
 RA Lee M.;
 RT "Cloning and characterisation of the human 5-HT_{2A} serotonin
 REBS receptor.";
 RL FEBS Lett. 355:242-246 (1994).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Brain;
 RA Puhl H.L. III, Ikeda S.R., Aronstam R.S.;
 RT "cDNA clones of human proteins involved in signal transduction
 RL Submitted (APR-2002) to the EMBL/GenBank/DBJ databases.
 CC -!- FUNCTION: This is one of the several different receptors for 5-
 CC as a neurotransmitter (serotonin), a biogenic hormone that functions
 CC this receptor is mediated by G proteins.
 CC -!- SUBCELLULAR LOCATION: Integral membrane protein.
 CC -!- SIMILARITY: Belongs to family 1 of G-protein coupled receptors.
 CC STRONGEST TO THE OTHER 5HT-5 SUBTYPE RECEPTORS.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL: X81411; CAA57168.1;
 CC EMBL: X81412; CAA57168.1; JOINED.
 CC EMBL: AF498985; AAM21132.1; --
 CC PIR: I37107; I37107.
 CC Genbank: HGNC:5300; HTR5A.
 CC MIM: 601305; --
 CC GO: GO:0005887; C: integral to plasma membrane; TAS.
 CC GO: GO:0004993; F: serotonin receptor activity; TAS.
 CC GO: GO:0007186; P: G-protein coupled receptor protein signalin. .; TAS.
 CC InterPro: IPR000276; GPCR_Rhodopsin.
 CC Pfam: PF00001; 7tm 1; 1.
 CC PRINTS: PR00237; GPCR_Rhodopsin.
 CC PROSITE: PS00237; G-PROTEIN RECEPTOR FL1; 1.
 CC PROSITE: PS00262; G-PROTEIN RECEPTOR FL2; 1.
 CC PROSITE: PS00262; G-PROTEIN RECEPTOR FL2; 1.
 CC G-protein coupled receptor; Transmembrane; Glycoprotein;
 KW Multigene family.
 FT DOMAIN 1 40 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 41 63 1 (POTENTIAL).
 FT DOMAIN 64 78 CYTOPLASMIC (POTENTIAL).
 FT TRANSMEM 79 99 2 (POTENTIAL).
 FT DOMAIN 100 115 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 116 137 3 (POTENTIAL).
 FT DOMAIN 138 158 CYTOPLASMIC (POTENTIAL).
 FT TRANSMEM 159 181 4 (POTENTIAL).
 FT DOMAIN 182 198 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 199 219 5 (POTENTIAL).
 FT DOMAIN 220 282 CYTOPLASMIC (POTENTIAL).
 FT TRANSMEM 283 303 6 (POTENTIAL).
 FT DOMAIN 304 320 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 321 341 7 (POTENTIAL).
 FT DOMAIN 342 357 CYTOPLASMIC (POTENTIAL).
 FT CARBOHYD 6 6 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 21 21 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT DISULFID 120 192 BY SIMILARITY.
 SQ SEQUENCE 357 AA; 40255 MW; 92F8A78C69169790 CRC64;
 Alignment Scores:
 Pred. No.: 2,19e-47 Length: 357
 Score: 1063.50 Matches: 226
 Percent Similarity: 72.98% Conservative: 36

Best Local Similarity: 62.95% Mismatches: 86
 Query Match: 49.37% Indels: 11
 DB: 1 Gaps: 6
 US-09-976-782-15 (1-1152) x SH5A_HUMAN (1-357)
 QY 65 ACCAGCAGCGGACCCCGGAGCCCAAGCCGAGAGGAGATACCTCGTTTCGACCCCGAGCGGC 124
 Db |||||
 Db 8 ThrSerPheSerLeuSerThrProSerProLeuGluThrAsnHisSerLeuGlyLysAsp 27
 QY 125 GCCGTCCTCGCGGCGGAGGCGGCGCTCTCTGCTTTCACGGTCTCTGTGTGAGCGTGG 184
 Db ::|||
 Db 28 AspLeuArgProSer---SerProLeuLeuSerValPheGlyValLeuLeuLeuThrLeu 46
 QY 185 CTAGTGTCTGCTGATCGTCCGCTTTCCTGTGGAACCTGCTGCTCGCTCGCTCCGCTCCG 244
 Db |||||
 Db 47 LeuGlyPheLeuValAlaAlaThrPheAlaThrPheLeuLeuValLeuAlaThrLeuLeu 66
 QY 245 CGGTCCTCGCTCCACCGCGCTGCGGATTAATCTGTGCTCGCTCGGCGCGCTCTCGGAC 304
 Db |||||
 Db 67 ArgValArgThrPheHisArgValProHisAsnLeuValAlaSerMetAlaValSerAsp 86
 QY 305 GAACCTAGTGGAGCGCTGCGGATGACCGAGCTGCGGAGTGTGAGTGTGCGAGCGGCGA 364
 Db |||||
 Db 87 ValLeuValAlaAlaLeuValMetProLeuSerLeuValHisGluLeuSer---GlyArg 105
 QY 365 CGTGGCTGCTGCGGCGGAGCTGTCACAGTGTGATCTCTTCGACCGCGGAGCTGT 424
 Db |||||
 Db 106 ArgTrpGlnLeuGlyArgLeuGlyGlnLeuTrpLeuAlaCysAspVal----- 122
 QY 425 CTGTGCTGCTGCGGCGGCTCGGGAACGTGGGCGGCTCGCTCGGCGCGGAGCGGCGC 484
 Db |||||
 Db 123 LeuCysCysThrAlaSerileTrpAsnValThrAlaLeuAlaLeuAspArgTyrTrpSer 142
 QY 485 ATCAGCAGCGCCTGACGACAGCTGCGCACCGGAGCGGCGCTCTGTGTGCTCATGATC 544
 Db |||||
 Db 143 IleThrArgHisMetGluTyrThrLeuArgThrArgLysCysValSerAsnValMetIle 162
 QY 545 GCGCTGCGGCGGCTGCGGCTGCGGCTGCGGCTGCGGCTGCGGCTGCGGCTGCGGCTG 604
 Db |||||
 Db 163 AlaLeuThrTrpAlaLeuSerAlaValIleSerLeuAlaProLeuLeuPheGlyTrpGly 182
 QY 605 GAGGTGTGCGAGCTGCGGCTCGGCTCGGCTCGGCTCGGCTCGGCTCGGCTCGGCTCG 664
 Db |||||
 Db 183 GluThrTyrSerGluGlySerGluGluCysGlnValSerArgGluProSerTyrAlaVal 202
 QY 665 TTCTCCACCGCGGCGCTTCCACCTGCGCTGCGGCTGCGGCTGCGGCTGCGGCTGCGG 724
 Db |||||
 Db 203 PheSerThrValGlyAlaPheTyrLeuProLeuCysValValLeuPheValTyrTrpLys 222
 QY 725 ATCTACGAGGCGGCAAGTTTCGTTTCGCG---CGCGCGGAGAGCTGCTGCGGTTG 781
 Db |||||
 Db 223 IleTyrLysAlaAlaLysPheArgValGlySerArgLysThrAsnSerValSerProIle 242
 QY 782 CCGGCCACCATGACGCTGAGTCCCAAGGTAAAGGACACCTGATGAGCTGAGCTGAGTGG 841
 Db ::|||
 Db 243 SerGluAlaValGluValLysAspSerAlaLys-----GlnProGlnMetVal 258
 QY 842 TTCACGGCACATTGCAACGACGCTGCTTCGAGGTGAGCGGAGCTCTCGGCGGAG 901
 Db |||||
 Db 259 PheThrValArg---HisAlaThrValThrPheGlnProGluGlyAspThrTrpArgGlu 277
 QY 902 CAGAAGGAGAGCGGAGCAGCATGATGTTGGGAAATCTGATTGGCGTGTTCGTGCTGTC 961
 Db |||||
 Db 278 GlnLysGluGlnArgAlaAlaLeuMetValGlyIleLeuLeuGlyValPheValLeuCys 297
 QY 962 TGGATCCCTCTCTTCCTGACGGAACCTATCAGCCCACTGTGTGCTGCGCTGCGCTGCG 1021
 Db |||||
 Db 298 TrpIleProPhePheLeuThrGluLeuIleSerProLeuCysSerCysAspIleProAla 317
 QY 1022 ATCTCGMAAAGCATATTTCTGTGGCTTGGCTACTCTCAATCTCTTCTCAACCCCTGATT 1081
 Db |||||
 Db 318 IleTrpLysSerIlePheLeuTrpLeuGlyTyrSerAsnSerPhePheAsnProLeuIle 337

QY 1082 TACACAGCTTTTACAGAACTACAAATGCTTCAGAGCCTCTTACTAAGCAG 1138
 Db 338 TyrThrAlaPheAsnLysAsnTyrAsnSerAlaPheLysAsnPheSerArgGln 356

RESULT 4

5H5A_MOUSE
 ID 5H5A_MOUSE STANDARD; PRT; 357 AA.
 AC P30966;
 DT 01-JUL-1993 (Rel. 26, Created)
 DT 01-JUL-1993 (Rel. 26, Last sequence update)
 DT 15-JUL-1998 (Rel. 36, Last annotation update)
 DE 5-hydroxytryptamine 5A receptor (5-HT-5A) (Serotonin receptor)
 DE (5-HT-5).
 GN HTR5A OR 5HT5A.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Brain;
 RC MEDLINE=93099851; PubMed=4464308;
 RA Plassat J.-L., Boscher U., Anlaik N., Hen R.;
 RA "The mouse 5HT5 receptor reveals a remarkable heterogeneity within
 RT the 5HT1D receptor family";
 RL EMO J. 11:4779-4786 (1992).
 RN [2]
 RP CHARACTERIZATION.
 RC TISSUE=Brain;
 RC MEDLINE=93196607; PubMed=9450829;
 RA Matthes H., Boscher U., Anlaik N., Grailhe R., Plassat J.-L.,
 RA Muscatelli F., Mattei M.-G., Hen R.;
 RT "Mouse 5-hydroxytryptamine 5A and 5-hydroxytryptamine 5B receptors
 RT define a new family of serotonin receptors: cloning, functional
 RT expression, and chromosomal localization.";
 RL Mol. Pharmacol. 43:313-319 (1993).
 CC -!- FUNCTION: This is one of the several different receptors for 5-
 CC hydroxytryptamine (serotonin), a biogenic hormone that functions
 CC as a neurotransmitter, a hormone, and a mitogen. The activity of
 CC this receptor is mediated by G proteins.
 CC -!- SUBCELLULAR LOCATION: Integral membrane protein.
 CC -!- TISSUE SPECIFICITY: Expressed predominantly in the central nervous
 CC system; in the cerebral cortex, hippocampus, habenula, olfactory
 CC bulb and granular layer of the cerebellum.
 CC -!- SIMILARITY: Belongs to family 1 of G-protein coupled receptors.
 CC STRONGEST TO OTHER 5HT-5 SUBTYPE RECEPTORS.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; Z18278; CAA79155.1; -
 CC MGD; MGI:96283; Htr5a.
 CC InterPro; IPR000276; GPCR_Rhodpsn.
 CC Pfam; PF00001; 7tm1.1;
 CC PRINTS; PR00237; GPCRHHODPSN.
 CC PROSITE; PS00237; G_PROTEIN_RECEPTOR_F1_1; 1.
 CC PROSITE; PS02662; G_PROTEIN_RECEPTOR_F2_1; 1.
 CC G-protein coupled receptor; Transmembrane; Glycoprotein;
 CC Multigene family.
 CC DOMAIN 1 40 EXTRACELLULAR (POTENTIAL).
 CC TRANSMEM 41 63 1 (POTENTIAL).
 CC DOMAIN 64 78 2 (POTENTIAL).
 CC TRANSMEM 79 99 3 (POTENTIAL).
 CC DOMAIN 100 115 3 (POTENTIAL).
 CC TRANSMEM 116 137 3 (POTENTIAL).
 CC DOMAIN 138 158 4 (POTENTIAL).
 CC TRANSMEM 159 181 4 (POTENTIAL).
 CC

FT DOMAIN 182 198 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 199 219 5 (POTENTIAL).
 FT DOMAIN 220 282 CYTOPLASMIC (POTENTIAL).
 FT TRANSMEM 283 303 6 (POTENTIAL).
 FT DOMAIN 304 320 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 321 341 7 (POTENTIAL).
 FT DOMAIN 342 357 CYTOPLASMIC (POTENTIAL).
 FT CARBOHYD 6 6 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 21 21 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT DISULFID 120 192 BY SIMILARITY.
 SQ SEQUENCE 357 AA; 40804 MW; 5F5D856AC477BFAC CRC64;
 Alignment Scores:
 Pred. No.: 2,67e-45 Length: 357
 Score: 1022.50 Matches: 219
 Percent Similarity: 71.11% Conservative: 37
 Best Local Similarity: 60.83% Mismatches: 93
 Query Match: 47.47% Indels: 11
 DB: 1 Gaps: 6
 US-09-976-782-15 (1-1152) x 5H5A_MOUSE (1-357)
 QY 65 ACCAGACGGAGACCCGGAGACCCCAAGCCGAGAGGATATCTGGTTCACCCCGAGCGGC 124
 Db 8 ThrSerPheSerLeuSerThrProSerSerLeuGluProAsnArgSerLeuAspThrGlu 27
 QY 125 GCCGTCCTGCGGCGGAGGCGCCCTCTCTCTCTCTACGGTCTCGGTGGTGCAGCGTG 184
 Db 28 ValLeuArgProSerArg---ProPheLeuSerAlaPheArgValLeuValLeuThrLeu 46
 QY 185 CTAGTGTCTGATCGCTGCCACCTTCTGTGGAACCTGCTGGTTCGGTCCACCATCCGCG 244
 Db 47 LeuGlyPheLeuAlaAlaAlaThrPheThrTrpAsnLeuValLeuAlaThrLeuLeu 66
 QY 245 CGGTGCTGCTGCTTCCACCGCGTCCCGGATAACTGTGTGGCTCGACGGCGCTCGGAC 304
 Db 67 LysValArgThrPheHisArgValProHisAsnLeuValAlaSerMetAlaIleSerAsp 86
 QY 305 GAAC TAGTGGAGCGCTGGCGATGCCACCGACCTGGCGAGTGGAGTGTTCACCGGGGA 364
 Db 87 ValLeuValAlaValLeuValMetProLeuSerLeuValHisGluLeuSer---GlyArg 105
 QY 365 CGTGGCTGCTGGCGCGGAGCTGTGCCAGCTGTGGATCTCTCTCGACGCCGAGCGCTGT 424
 Db 106 ArgTrpGlnLeuGlyArgLeuLeuGlnLeuTrpIleAlaCysAspVal-----122
 QY 425 CTGTGCTGCCCGCGGCTCGGAAACGTGGCGCCATCGCCCTGGCGCGGAGCGGCGCC 484
 Db 123 LeuCysThrAlaSerIleTrpAsnValThrAlaIleAlaLeuAspArgTyrTrpSer 142
 QY 485 ATCACGGGACCTGCAGCACACCTGCGACCGCCGCGCGCGCTCGTTCATGATC 544
 Db 143 IleThrArgHisLeuGluTyrThrLeuArgThrArgLysArgValSerAsnValMetIle 162
 QY 545 CGCTGCGCGCGGTGCGCTCATCCCTCGCGCGCTGCTCTTTGGCGGGCG 604
 Db 163 LeuLeuThrTrpAlaLeuSerThrValIleSerLeuAlaProLeuLeuPheGlyTrpGly 182
 QY 605 GAGTGTGCGAGCTCGGCTCCAGCTGCCAGTGCAGTGCAGCGCGGACCCCTCTATGCCGCC 664
 Db 183 GluThrTyrSerGluProSerGluCysGlnValSerArgGluProSerTyrThrVal 202
 QY 665 TTCTCCACCGCGGCGCTTCCACCTGCGGCTGGCGTGGTGGCTGTGTACCGGAG 724
 Db 203 PheSerThrValGlyAlaPheTyrLeuProLeuTrpLeuValLeuPheValTyrTrpLys 222
 QY 725 ATCTACGAGGCGCAAGTTTCTGGC---CGCGCGGAGAGCTGTGCTCGCGTGTG 781
 Db 223 IleThrArgAlaAlaLysPheArgMetGlySerArgLysThrAsnSerValSerProVal 242
 QY 782 CCGGCCACCATGCAGGTGAGGTCCCAAGTAAGGAAGCACCTGATGAGGCGTGAAGTGTGTG 841
 Db 243 ProGluAlaValGlu-----ValLysAsnAlaThrGlnHisProGlnMetVal 258


```

Db 228 LysPheArgMetGlySerArgGlyThrAsnSerValSerProGluAlaValGlu 247
Qy 797 GTGAGGTCGAAGTAAGGAAGACCTGATGAGGCTGAAGTGGTGTTCACGGCACATTGC 856
Db 248 -----VallysAspAlaSerGlnHisProGlnMetValPheThrValArg--- 262
Qy 857 AAAGCAACGGTGCTCTCCAGGTGAGCGGCGACTCTCGCGGAGCGAGAGGAGGCGGA 916
Db 263 HisAlaThrValThrPheGlnThrGluGlyAspThrTrpArgGluGlnLysGluGlnArg 282
Qy 917 GCAGCCATGATGGTGGGAATCTCATGTCGCGCTGTTGCTGCTGTCGATCCCTTCCTTC 976
Db 283 AlaAlaLeuMetValGlyIleLeuIleGlyValPheValLeuCysTrpPheProPhePhe 302
Qy 977 CTGACGGAACATCATCAGCCCACTCTGTCGCTGCGAGCTGCGCCCACTGGAAGAAGATA 1036
Db 303 ValThrGluLeuIleSerProLeuCysSerTrpAspIleProAlaLeuTrpLysSerIle 322
Qy 1037 TTTCTGTGGCTTGGCTACTCCAAATCTCTTCTCAACCCCTGATTACACAGCTTTTAAC 1096
Db 323 PheLeuTrpLeuGlyTrpSerAsnSerPhePheAsnProLeuIleTrpThrAlaPheAsn 342
Qy 1097 AAGAACTACAACATGCTTCAAGAGCCTCTTTACTAGCAGAGA 1141
Db 343 ArgSerTrpSerAlaPheLysValPhePheSerLysGlnGln 357

RESULT 6
SH1B HUMAN STANDARD; PRT; 390 AA.
AC 28222;
DT 01-DEC-1992 (Rel. 24, Created)
DT 01-DEC-1992 (Rel. 24, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE 5-hydroxytryptamine 1B receptor (5-HT-1B) (Serotonin receptor)
DE (5-HT-1D-beta) (Serotonin 1D beta receptor) (S12).
GN HTR1B OR HTR1DB.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
OX NCBI_TaxID=9606;
RN 1;
RS SEQUENCE FROM N.A.
RC TISSUE=Placenta;
RX MEDLINE=92228840; PubMed=1565658;
RA Weinshank R.L., Zgombick J.M., Macchi M.J., Branchek T.A.,
RA Hartig P.R.;
RT "Human serotonin 1D receptor is encoded by a subfamily of two
RT distinct genes: 5-HT1D alpha and 5-HT1D beta.";
RL Proc. Natl. Acad. Sci. U.S.A. 89:3630-3634(1992).
RN 2;
RS SEQUENCE FROM N.A.
RX MEDLINE=92302275; PubMed=1351684;
RA Demchyshyn L., Sunahara R.K., Miller K., Teitler M., Hoffman B.J.,
RA Kennedy J.L., Seeman P., van Tol H.H.M., Niznik H.B.;
RT "A human serotonin 1D receptor variant (5HT1D beta) encoded by an
RT intronless gene on chromosome 6.";
RL Proc. Natl. Acad. Sci. U.S.A. 89:5522-5526(1992).
RN 3;
RS SEQUENCE FROM N.A.
RX MEDLINE=92304270; PubMed=1610347;
RA Mochizuki D., Yuyama Y., Tsujita R., Komaki H., Sagai H.;
RT "Cloning and expression of the human 5-HT1B-type receptor gene.";
RL Biochem. Biophys. Res. Commun. 185:517-523(1992).
RN 4;
RS SEQUENCE FROM N.A.
RX MEDLINE=92246962; PubMed=1315531;
RA Hamblin M.W., Metcalf M.A., McGuffin R.W., Karpellis S.;
RT "Molecular cloning and functional characterization of a human 5-HT1B
RT serotonin receptor: a homologue of the rat 5-HT1B receptor with
RT 5-HT1D-like pharmacological specificity.";
RL Biochem. Biophys. Res. Commun. 184:752-759(1992).
RN 5;

```

```

RP SEQUENCE FROM N.A.
RX MEDLINE=92218412; PubMed=1559993;
RA Levy F.O., Gudermann T., Perez-Reyes E., Birnbaumer M., Kaumann A.J.,
RA Birnbaumer L.;
RT "Molecular cloning of a human serotonin receptor (S12) with a
RT pharmacological profile resembling that of the 5-HT1D subtype.";
RL J. Biol. Chem. 267:7553-7562(1992).
RN 6;
RS SEQUENCE FROM N.A.
RX MEDLINE=93024334; PubMed=1328844;
RA Veldman S.A., Bienkowski M.J.;
RT "Cloning and pharmacological characterization of a novel human 5-
RT hydroxytryptamine1D receptor subtype.";
RL Mol. Pharmacol. 42:439-444(1992).
RN 7;
RS SEQUENCE FROM N.A.
RX MEDLINE=92210526; PubMed=1348246;
RA Jin H., Oksenberg D., Ashkenazi A., Peroutka S.J., Duncan A.M.V.,
RA Rozmahel R., Yang Y., Mengod G., Palacios J.M., O'Dowd B.F.;
RT "Characterization of the human 5-hydroxytryptamine1B receptor.";
RL J. Biol. Chem. 267:5735-5738(1992).
RN 8;
RS SEQUENCE FROM N.A.
RX Kitano T., Kobayakawa H., Saitou N.;
RT "Silver project.";
RL Submitted (APR-2000) to the EMBL/GenBank/DBJ databases.
RN 9;
RS SEQUENCE FROM N.A.
RX Kopatz S.A., Aronstam R.S., Sharma S.V.;
RT "cDNA clones of human proteins involved in signal transduction
RT sequenced by the Guthrie cDNA resource center (www.cdna.org).";
RL Submitted (JAN-2003) to the EMBL/GenBank/DBJ databases.
RN 10;
RS SEQUENCE FROM N.A.
RX Westharp J.;
RT Submitted (JUL-1999) to the EMBL/GenBank/DBJ databases.
RN 11;
RS PALMITOYLATION, AND PHOSPHORYLATION.
RX MEDLINE=94032297; PubMed=8218242;
RA Ng G.Y.K., George S.R., Zastawny R.L., Caron M., Bouvier M.,
RA Dennis M., O'Dowd B.F.;
RT "Human serotonin1B receptor expression in Sf9 cells: phosphorylation,
RT palmitoylation, and adenylyl cyclase inhibition.";
RL Biochemistry 32:11727-11733(1993).
RN 12;
RS VARIANT CYS-124.
RX MEDLINE=95100945; PubMed=7802650;
RA Nothen M.M., Erdmann J., Shimron-Abarbanell D., Propping P.;
RT "Identification of genetic variation in the human serotonin 1D beta
RT receptor gene.";
RL Biochem. Biophys. Res. Commun. 205:1194-1200(1994).
RN 13;
RS FUNCTION: This is one of the several different receptors for 5-
RX hydroxytryptamine (serotonin), a biogenic hormone that functions
RX as a neurotransmitter, a hormone, and a mitogen. The activity of
RX this receptor is mediated by G proteins that inhibit adenylylate
RX cyclase activity.
RN 14;
RS SUBCELLULAR LOCATION: Integral membrane protein.
RX Phosphorylated, and palmitoylated. Desensitization of the
RX receptor may be mediated by its phosphorylation.
RN 15;
RS SIMILARITY: Belongs to family 1 of G-protein coupled receptors.
RX Strongest to the other 5HT-1 subtype receptors.
RN 16;
RS This SWISS-PROT entry is copyright. It is produced through a collaboration
RX between the Swiss Institute of Bioinformatics and the EMBL outstation -
RX the European Bioinformatics Institute. There are no restrictions on its
RX use by non-profit institutions as long as its content is in no way
RX modified and this statement is not removed. Usage by and for commercial
RX entities requires a license agreement (See http://www.isb-sib.ch/announce/
RX or send an email to license@isb-sib.ch).
RN 17;
RX EMBL; D10995; BAA01763.1; -
RX EMBL; M81590; AAA60316.1; -
RX EMBL; M75128; AAA58675.1; -

```

DR EMBL; M89478; -, NOT ANNOTATED_CDS.
 DR EMBL; L09732; AAA36030.1; -
 DR EMBL; M83180; AAA36029.1; -
 DR EMBL; AB041370; BAA94455.1; -
 DR EMBL; AY25227; AAO67712.1; -
 DR EMBL; AL049595; CAB51537.1; -
 DR PIR; JN0268; JN0268.
 DR Genew; HGNC:5287; HTRIB.
 DR MIM; 182131; -
 DR GO; GO:0005887; C:integral to plasma membrane; TAS.
 DR GO; GO:0004993; F:serotonin receptor activity; TAS.
 DR GO; GO:0007187; P:G-protein signaling, coupled to cyclic nucl. .; TAS.
 DR GO; GO:0007268; P:synaptic transmission; TAS.
 DR InterPro; IPR000276; GPCR_Rhodopsn.
 DR Pfam; PF00001; 7tm.1; 1
 DR PRINTS; PRO0237; GPCRHOPOPSN.
 DR PROSITE; PS00237; G PROTEIN RECP Fl_1; 1.
 DR PROSITE; PS00262; G PROTEIN RECP Fl_2; 1.
 DR G-protein coupled receptor; Transmembrane; Glycoprotein;
 KW Multigene family; Lipoprotein; Palmitate; Phosphorylation;
 KW Polymorphism.
 FT DOMAIN 1 49 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 50 73 1 (POTENTIAL).
 FT DOMAIN 74 86 CYTOPLASMIC (POTENTIAL).
 FT TRANSMEM 87 109 2 (POTENTIAL).
 FT DOMAIN 110 119 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 120 145 3 (POTENTIAL).
 FT DOMAIN 146 165 CYTOPLASMIC (POTENTIAL).
 FT TRANSMEM 166 187 4 (POTENTIAL).
 FT DOMAIN 188 205 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 206 229 5 (POTENTIAL).
 FT DOMAIN 230 315 CYTOPLASMIC (POTENTIAL).
 FT TRANSMEM 316 340 6 (POTENTIAL).
 FT DOMAIN 341 347 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 348 373 7 (POTENTIAL).
 FT DOMAIN 374 390 CYTOPLASMIC (POTENTIAL).
 FT CARBOHYD 24 24 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 32 32 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT DISULFID 122 199 BY SIMILARITY.
 FT LIPID 388 388 S-palmitoyl cysteine (Potential).
 FT VARIANT 124 124 F -> C (in dbSNP:130060).
 FT VARIANT 219 219 F -> L (in dbSNP:130061).
 FT VARIANT 367 367 I -> V (in dbSNP:130063).
 FT VARIANT 374 374 E -> K (in dbSNP:130064).
 FT SEQUENCE 390 AA; 43568 MW; CD874DC7EB44CF12 CRC64;
 Alignment Scores:
 Pred. No.: 4 31e-16 Length: 390
 Score: 448.00 Matches: 127
 Percent Similarity: 48.35% Conservative: 63
 Best Local Similarity: 32.32% Mismatches: 147
 Query Match: 20.80% Indels: 56
 DB: 1 Gaps: 13
 US-09-976-782-15 (1-1152) x SH1B_HUMAN (1-390)
 QY 59 CCGAGACCGAGCGGACCCGGGACCCCAAGCCGAGAGGGATACGCGTTGCGACCCG 118
 Db 10 ProProProAlaGlySerGluThrTrpValProGlnAlaAsnSerSerAlaPro 29
 QY 119 AGC-----GGCGCGCTCTCGCGGCCGA 142
 Db 30 SerGlnAsnCysSerAlaLysAspTyrIleTyrGlnAspSerIleSerLeuPro----- 47
 QY 143 GGGCGCGCCCTCTCTCTCTACGGTCCTGGTGCGACGCTGCTAGTGTGCTGTCGCT 202
 Db 48 -----TrpLysValLeuLeuValMetLeuLeuAlaLeuIleThrLeu 61
 QY 203 GCACCTTCTGTGGAACTGCTGTGCTGCGTCCACCATCCCGGGGTCCGCTTCCAC 262

RESULT 7

SH1B_PANTR

ID SH1B_PANTR STANDARD; PRT; 390 AA.

AC P6020;

DT 15-MAR-2004 (Rel. 43, Created)

DT 15-MAR-2004 (Rel. 43, Last sequence update)

Db 62 AlaThrThrLeuSerAsnAlaPheValIleAlaThrValTyrArgThrArgLysLeuHis 81
 QY 263 CGCGTCCGCATAACTTGTGGCTCGACGGCGCTCGACGAAGTCTCGAGCGAGCGCTG 322
 Db 82 ThrProAlaAsnTyrLeuIleAlaSerLeuAlaValThrAspLeuLeuValSerIleLeu 101
 QY 323 CGGATGCCACCGAGCTGGCGAGTGTGCGACGGCGGCGAGCTGGCTGGTGGCGCGG 382
 Db 102 ValMetPro-----IleSerThrMetTyrThrValThrGlyArgTrpThrLeuGlyGln 119
 QY 383 AGCTGTGCGACGTGTGATCTCTTCGACGGCGGAGCCGTGTCTGTCTGCCCGCGCGC 442
 Db 120 ValValCysAspPheTrpLeuSerSerAsp-----IleThrCysCysThrAlaSer 136
 QY 443 CTCGGGAACGTGGCGCCATCGCCCTGGCGGGCGGAGGGGCGCATCACAGGCGACGTGCG 502
 Db 137 IleLeuHisLeuCysValIleAlaLeuAspArgTyrTrpAlaIleThrAspAlaValGlu 156
 QY 503 CACAGCTGCGCACCGCGAGCGCGCTGTCTCATGATCGCGCTCGCGCGGGTGGCGG 562
 Db 157 TyrSerAlaLysArgThrProLysArgAlaAlaValMetIleAlaLeuValTrpValPhe 176
 QY 563 TCGGCGCTCATCCCTCGCGCGCTGTCTTTGGCGGGCGAGGTGTCGAGCGTCCG 622
 Db 177 SerIleSerIleSerLeuProPro---PhePheTrpArgGlnAlaLysAlaGluGlu 195
 QY 623 CTCGAGCGTGCAGGTGAGCGGGAGCC---TCCTATGCCGCTTCTCCACCGCGCGC 679
 Db 196 ValSerGluCysValValAsnThrAspHisIleLeuTyrThrValTyrSerThrValGly 215
 QY 680 GCCTTCCACCTCGCGCTGGCGTGTGTCTACCGGAAGATCTACGAGGCGGCGC 739
 Db 216 AlaPheTyrPheProThrLeuLeuIleAlaLeuTyrGlyArgIleTyrValGluAla 235
 QY 740 AAGTTTCGTTTC-----GGCGGCGCGC---CGGAGAGGTGTG 772
 Db 236 ArgSerArgIleLeuLysGlnThrProAsnArgThrGlyLysArgLeuThrArgAlaGln 255
 QY 773 CTG-----CCGTGCGCGCCACCATCGAGTGTGAGGTCCAGGTAAAG 814
 Db 256 LeuIleThrAspSerProGlySerThrSerSerValThrSerIleAsnSerArgValPro 275
 QY 815 GAAGCACCTGATGAGGT---GAAGTGTGTTCACGGCACATTGCAAGCAACGCGTGTCC 871
 Db 276 AspValProSerGluSerGlySerProValTyrValAsnGlnValLysValArgValSer 295
 QY 872 TTCCAGGTGAGCGGGACTCTCTGGCGGGGAGCGAAG-----GAGAGG 913
 Db 296 -----AspAlaLeuLeuGluLysLysLysLeuMetAlaAlaArgGluArg 310
 QY 914 CGAGCAGCATGATGTTGGGAATTCGATTGGCGTGTGTGTCTGTGCTGGATCCCTTC 973
 Db 311 LysAlaThrLysThrLeuGlyIleIleLeuGlyAlaPheIleValCysTrpLeuProPhe 330
 QY 974 TTCCTGACGAACTCATCAGCCACTCTGT-----GCCTGCGAGCTGCGCCCATCTGG 1027
 Db 331 PheIleIleSerLeuValMetProIleCysLysAspAlaCysTrpPheHisLeuAlaIle 350
 QY 1028 AAAAGCATTTCTGTGGCTTGGCTTACTCCAAATCTTTCTTCAACCCCTGATTCACA 1087
 Db 351 PheAspPhePheThrTrpLeuGlyTyrLeuAsnSerLeuIleAsnProIleIleTyrThr 370
 QY 1088 GCTTTTAAACAAGAACTACAAACATGCTTCAAGAGCCTC 1126
 Db 371 MetSerAsnGluAspPheLysGlnAlaPheHisLysLeu 383

15-MAR-2004 (Rel. 43, Last annotation update)
 5-hydroxytryptamine 1B receptor (5-HT-1B) (Serotonin receptor).
 HTR1B.
 OS Pan troglodytes (Chimpanzee).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Pan.
 OK NCBI_TaxID=9598;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Kitano T., Kobayakawa H., Saitou N.;
 RT "Silver project";
 RL Submitted (APR-2000) to the EMBL/GenBank/DBJ databases.
 CC -!- FUNCTION: This is one of the several different receptors for 5-
 hydroxytryptamine (serotonin), a biogenic hormone that functions
 as a neurotransmitter, a hormone, and a mitogen. The activity of
 this receptor is mediated by G proteins that inhibit adenylyate
 cyclase activity (By similarity).
 CC -!- SUBCELLULAR LOCATION: Integral membrane protein.
 CC -!- PTM: Phosphorylated, and palmitoylated. Desensitization of the
 receptor may be mediated by its phosphorylation (By similarity).
 CC -!- SIMILARITY: Belongs to family 1 of G-protein coupled receptors.
 CC Strongest to the other 5HT-1 subtype receptors.

 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
 or send an email to license@isb-sib.ch).

 CC EMBL: AB041371; BAA94456.1;
 DR PROSITE; PS00237; G-PROTEIN_RESEP_F1_1; 1.
 DR PROSITE; PS00262; G-PROTEIN_RESEP_F1_2; 1.
 KW G-protein coupled receptor; Transmembrane; Glycoprotein;
 KW Multigene family; Lipoprotein; Palmitate; Phosphorylation.
 FT DOMAIN 1 49
 FT TRANSMEM 50 73
 FT DOMAIN 74 86
 FT TRANSMEM 87 109
 FT DOMAIN 110 119
 FT TRANSMEM 120 145
 FT DOMAIN 146 165
 FT TRANSMEM 166 187
 FT DOMAIN 188 205
 FT TRANSMEM 206 229
 FT DOMAIN 230 315
 FT TRANSMEM 316 340
 FT DOMAIN 341 347
 FT TRANSMEM 348 373
 FT DOMAIN 374 390
 FT CARBOHYD 24 24
 FT CARBOHYD 32 32
 FT DISULFID 122 199
 FT LIPID 388 388
 SQ SEQUENCE 390 AA; 43568 MW; CD874DC7EB44CF12 CRC64;

 Alignment Scores:
 Pred. No.: 4.31e-16 Length: 390
 Score: 448.00 Matches: 127
 Percent Similarity: 48.35% Conservative: 63
 Best Local Similarity: 32.32% Mismatches: 147
 Query Match: 20.80% Indels: 56
 DB: 1 Gaps: 13

 US-09-976-782-15 (1-1152) x 5H1B_PANTR (1-390)
 QY 59 CCCGAGACACGACGACCGGACCCCAAGCCGAGAGGATACGCGTTCGACCCG 118
 Db 10 ProProProAlaGlySerGluThrTrpValProGlnAlaAsnLeuSerSerAlaPro 29
 QY 119 AGC-----GGCGCGTCTCTCCCGGCGCA 142
 |||
 |||

30 SerGlnAsnCysSerAlaLysAspTyrIleTyrGlnAspSerIleSerLeuPro----- 47
 143 GGGCGGCGCTTCTCTGCTTTCACGCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 202
 Db :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||: 61
 48 -----TrpLysValLeuLeuValMetLeuLeuAlaLeuIleThrLeu 61
 QY :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||: 262
 203 GCCACCTTCTGCTGGAACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 262
 Db :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||: 81
 62 AlaThrThrLeuSerAsnAlaPheValIleAlaThrValTyrArgThrArgLysLeuHis 81
 QY :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||: 322
 263 CGCGTGGCGCATAACTTGTGTGCTGCGAGCGCTCTCGGACGAACTAGTGGCGAGCGTGT 322
 Db :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||: 101
 82 ThrProAlaAsnTyrLeuIleAlaSerLeuAlaValThrAspLeuLeuValSerIleLeu 101
 QY :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||: 382
 323 CGCATGCCACCGAGCTGCGAGTGTGCTGCGAGCGGCGGCGAGCTGCGGCTGCTGCGGCGG 382
 Db :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||: 119
 102 ValMetPro-----IleSerThrMetTyrThrValThrGlyArgTyrThrLeuGlyGln 119
 QY :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||: 442
 383 AGCCTGTGTCACGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 442
 Db :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||: 136
 120 ValValCysAspPheTrpLeuSerSerAsp-----IleThrCysCysThrAlaSer 136
 QY :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||: 502
 443 CTCGGGAACGTGGCGGCATGCGCTGCGCGCGAGCGGCGGCATCACACGCGACCTGCGAG 502
 Db :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||: 156
 137 IleLeuHisLeuCysValIleAlaLeuAspArgTyrTrpAlaIleThrAspAlaValGlu 156
 QY :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||: 562
 503 CACAGCTGCGCACCAGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 562
 Db :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||: 176
 157 TyrSerAlaLysArgThrProLysArgAlaAlaValMetIleAlaLeuValTrpValPhe 176
 QY :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||: 622
 563 TCGCGCTCATCGCTCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 622
 Db :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||: 195
 177 SerIleSerIleSerLeuProPro---PhePheTrpArgGlnAlaLysAlaGluGlu 195
 QY :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||: 679
 623 CTCAGCGCTGCCAGGTGAGCGCGGAACCC---TCCTATGCGCGCTTCTCCACCGCGGCG 679
 Db :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||: 215
 196 ValSerGluCysValValAsnThrAspHisIleLeuTyrThrValTyrSerThrValGly 215
 QY :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||: 739
 680 GCCTTCCACCTGCGCGTGGCGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 739
 Db :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||: 235
 216 AlaPheTyrPheProThrLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeu 235
 QY :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||: 772
 740 AAGTTTCGTTTC-----GGCGCGCGC---CGGAGAGCTGTG 772
 Db :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||: 255
 236 ArgSerArgIleLeuLysGlnThrProAsnArgThrGlyLysArgLeuThrArgAlaGln 255
 QY :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||: 814
 773 CTG-----CCGTGCGCGCCACCATGCGAGGTGAGGTGCGAGGTAAAG 814
 Db :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||: 275
 256 LeuIleThrAspSerProGlySerThrSerSerValThrSerIleAsnSerArgValPro 275
 QY :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||: 871
 815 GAACACCTGATGAGGCT---GAAGTGTGTTCACGGGCACATTGCAAGCAACGCTGTCC 871
 Db :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||: 295
 276 AspValProSerGluSerGlySerProValTyrValAsnGlnValLysValArgValSer 295
 QY :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||: 913
 872 TTCAGGTGAGCGGAGCTCTCGTGGCGGAGCAAG-----GAGAGG 913
 Db :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||: 310
 296 -----AspAlaLeuLeuLysLysLysLysLysLysLysLysLysLysLysLysLysLysLys 310
 QY :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||: 973
 914 CGAGCAGCCATGATGTTGGGGAATTCATGTTGGCGGTGTTGCTGCTGCTGCTGCTGCTGCTGCT 973
 Db :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||: 330
 311 LysAlaThrLysThrLeuGlyIleIleLeuGlyAlaPheIleValCysTrpLeuProPhe 330
 QY :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||: 1027
 974 TTCCTGACGNACTCATCAGCCCACTCTGT-----GCCTGACGCTGCCCCCATCTGG 1027
 Db :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||: 350
 331 PheIleIleSerLeuValMetProIleCysLysAspAlaCysTrpPheHisLeuAlaIle 350
 QY :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||: 1087
 1028 AAAAGCATATTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1087
 Db :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||: 370
 351 PheAspPhePheThrTrpLeuGlyTyrLeuAsnSerLeuLeuLeuLeuLeuLeuLeuLeuLeu 370
 QY :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||: 1126
 1088 GCTTTTAAAGAACTACAACTGCTTCAAGAGCCTC 1126
 Db :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||: 383
 371 MetSerAsnGluAspPheLysGlnAlaPheHisLysLeu 383

QY 1022 ATCTGGGAAGCATATTCTGCTGCTGCTACTCCAAATCTTTCTTCAACCCCTGATT 1081
Db 345 AlallePheaspPheAenrPheLeuglyTyriLeuasnSerLeulleAsnProille 364
QY 1082 TACACAGCTTTTAAACAGAACTACAAATCGCTTCAAGAGCCTC 1126
Db 365 TyrThrMetProasnGluaspPheLysGlnAlaPheHisLysLeu 379
RESULT 9
SHIB_CAVPO
ID SHIB_CAVPO STANDARD; PRT; 389 AA.
AC O08832;
DT 15-JUL-1998 (Rel. 36, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE 5-hydroxytryptamine 1B receptor (5-HT-1B) (Serotonin receptor)
DE (5-HT1B receptor).
GN HTR1B.
OS Cavia porcellus (Guinea pig).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Hystricognathi; Caviidae; Cavia.
OC NCBI_TaxID=10141;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=9736862; PubMed=9225276;
RA Zgombick J.M., Bard J.A., Kucharewicz S.A., Urquhart D.A.,
RA Weinschank R.L., Branchek T.A.;
RT "Molecular cloning and pharmacological characterization of guinea pig
RT 5-HT1B and 5-HT1D receptors";
RL Neuropharmacology 36:513-524 (1997).
CC -1- FUNCTION: This is one of the several different receptors for 5-
CC hydroxytryptamine (serotonin), a biogenic hormone that functions
CC as a neurotransmitter, a hormone, and a mitogen. The activity of
CC this receptor is mediated by G proteins that inhibit adenylyate
CC cyclase activity.
CC -1- SUBCELLULAR LOCATION: Integral membrane protein.
CC -1- SIMILARITY: Belongs to family 1 of G-protein coupled receptors.
CC Strongest to the other 5HT-1 subtype receptors.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL: U82175; AAB58500.1; -;
DR InterPro: IPR000276; GPCR_Rhodopsn.
DR Pfam: PF00061; 7tm_1; 1.
DR PRINTS; PR00237; GPCRHOPOPSN.
DR PROSITE; PS00237; G PROTEIN RECF F1 1; 1.
DR PROSITE; PS0262; G PROTEIN RECF F1 2; 1.
KW G-protein coupled receptor; Transmembrane; Glycoprotein;
KW Multigene family; Lipoprotein; Palmitate.
FT DOMAIN 1 50 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 51 74 1 (POTENTIAL).
FT DOMAIN 75 85 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 86 112 2 (POTENTIAL).
FT DOMAIN 113 122 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 123 144 3 (POTENTIAL).
FT DOMAIN 145 166 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 167 186 4 (POTENTIAL).
FT DOMAIN 187 207 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 208 227 5 (POTENTIAL).
FT DOMAIN 228 314 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 315 335 6 (POTENTIAL).
FT DOMAIN 336 349 7 (POTENTIAL).
FT TRANSMEM 350 372 CYTOPLASMIC (POTENTIAL).
FT DOMAIN 373 389 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 24 24

FT CARBOHYD 31 31 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT DISULFID 121 198 BY SIMILARITY.
FT LIPID 387 387 S-palmitoyl cysteine (Potential).
SQ SEQUENCE 389 AA; 43110 MW; C057CAB0A7FEE3C6 CRC64;
Alignment Scores:
Pred. No.: 6,49e-16 Length: 389
Score: 444.50 Matches: 125
Percent Similarity: 48.44% Conservative: 61
Best Local Similarity: 32.55% Mismatches: 159
Query Match: 20.64% Indels: 39
DB: 1 Gaps: 12
US-09-976-782-15 (1-1152) x SHIB_CAVPO (1-389)
QY 59 CCGAGACAGCAGCGGACCGGAGCCCAAGCCCGAGAGGATCTCGGTTCGACCCCG 118
Db 10 ProProAlaValLeuGlySerGlnThrGlyLeuProHisAlaAsnValSerAlaProPro 29
QY 119 AGCGGCGCGCTCTCGCGGCGCGGAGCGCGCTTCTCTGTC-----TTCACGCTC 169
Db 30 AsnAsnCysSerAlaProSerHisIleTyrGlnAspSerIleAlaLeuProTyrLysVal 49
QY 170 CTGCTGTGTGACGCTGTAGTCTGCTGATCGCTCCACTTCTCTGTGTGGAACCTGCTGTT 229
Db 50 LeuLeuValValLeuLeuAlaLeuLeuLeuAlaThrThrLeuSerAsnAlaPheVal 69
QY 230 CCGGTACCATCCCGCGGCTCGCTTCCACCGCTGCGGATACCTTGGTGGCGCTCG 289
Db 70 IleAlaThrValTyrArgThrArgLysLeuHisThrProAlaAsnTyrLeuIleAlaSer 89
QY 290 ACGGCGCTCTCGAGCAACTAGTGGCAGCTGCGATGCCACGCGCTCGGAGGTGAG 349
Db 90 LeuAlaPheThrAspLeuLeuValSerIleLeuValMetPro-----IleSerThrMet 107
QY 350 CTGTTCGACCGGCGACGCTCGCTGCTGGCGCGGAGCTGTGCCACGTGTGTGATCTCTTC 409
Db 108 TyrThrValThrGlyArgTyrThrLeuGlyGlnAlaLeuCysAspPheTyrLeuSerSer 127
QY 410 GAGCGCGGAGCTGTCTGTCTGCTGCCCGCGGCTCGGAGACGTGGCGGCATCGCCTG 469
Db 128 Asp-----IleThrCysThrAlaSerIleMetHisLeuCysValIleAlaLeu 144
QY 470 GCGCGACGCGGCGCATCACAGGCACCTGCAGCACACGCTGCGCACCCGCGAGCGGCC 529
Db 145 AspArgTyrTrpAlaIleThrAspAlaValGlyTyrSerAlaLysArgThrProArgArg 164
QY 530 TCGTGTCTCATGATCGCGCTCGCCCGGCTCGCTGCGGCTCATCGCCCTCGCGCGCTG 589
Db 165 AlaAlaGlyMetIleAlaLeuValTyrValPheSerIleCysIleSerLeuProPro--- 183
QY 590 CTCTTTGGCGGCGGAGGTGTGGAGCTGCGCTCGGCTCCAGCGCTCGAGGTGAGCGGGAA 649
Db 184 PhePheTyrArgGlnAlaLysAlaGluGluValLeuLeuAspCysLeuValAsnThrAsp 203
QY 650 ---CCCTCTATCGCGCTTCTCCACCGCGGCGCTTCCACCTGCGCTGGGTGGGTGGT 706
Db 204 HisValLeuTyrThrValTyrSerThrGlyGlyAlaPheTyrLeuProThrLeuLeuLeu 223
QY 707 CCGTTTGTCTACCGGAAGATCTACGAGCGCGCCAAAGTTTCGTTTC----- 751
Db 224 IleAlaLeuTyrGlyArgIleTyrValGluAlaArgSerArgIleLeuLysGlnThrPro 243
QY 752 -----GGCGCGCGC---CGAGAGCTGTGCTG-----CCGTTG 781
Db 244 AsnLysThrGlyLysArgLeuThrArgAlaGlnLeuIleThrAspSerProGlySerThr 263
QY 782 CCGGCCACCATGCGAGGTGAGTCCAAGGTAAAGGAGCACCT---GATGAGGCTGAAGT 838
Db 264 SerSerValThrSerIleAsnSerArgAlaProGluValProCysAspSerGlySerPro 283
QY 839 GTGTTCACGGCACATTCGAAAGCAACCGTGTCTTCCAGGTGAGCGGGGACTCTCTGCGG 898
|||||


```

Qy 467 CTGGCGCGCGAGCGGCATCACAGCGCACTGCAGCACACGCTGGCGCACCCGCGAGCGCG 526
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
145 LeuAspArgTyrTrpAlaIleThrAspAlaValGluTyrSerAlaLysArgThrProLys 164
Qy 527 GCCTCGTGTCTCARGATCGCGCTCGCGCGGTGGCGTGGCGCTCATGCCCTCGCGCGCG 586
Db ::::: ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
165 ArgAlaAlaIleMetIleArgLeuValTrpValPheSerIleCysIleSerLeuProPro 184
Qy 587 CTGCTCTTTGGCGCGGAGGTGTGGACGCTGGCTCGGCTCCAGCGCTGCACGTTGAGCGCG 646
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
185 ---PhePheTrpArgGlnAlaLysAlaGluGluValSerGluCysLeuValAsnThr 203
Qy 647 GAA---CCCTCTATGCGCGCTTCTCCACCGCGCGCTTCCACCTGCCCTGGCGTGG 703
Db ::::: ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
204 AspHisValLeuTyrThrValTyrSerThrValGlyAlaPheTyrLeuProThrLeuLeu 223
Qy 704 GTGCGCGTTGTCTACCGAAGATCTACGAGCGCGCAAGTTTCGTTC----- 751
Db ::::: ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
224 LeuIleAlaLeuTyrGlyArgIleTyrValGluAlaArgSerArgIleLeuLysGlnThr 243
Qy 752 -----GGCGCGCGC---CGGAGAGCTGTGCTG-----CCG 778
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
244 ProAsnArgThrGlyLysArgLeuThrArgAlaGlnLeuIleThrAspSerProGlySer 263
Qy 779 TTGCGGCGCCACATGCAGTGCAGGTGCCAAGGTAAAGGAAGCACCTGATGAGCT---GAA 835
Db ::::: ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
264 ThrThrSerValThrSerIleAsnSerArgAlaProAspValProSerGluSerGlySer 283
Qy 836 GTGCGTGTACCGGCACATGTCGAGCAAGCGGTGCTTCCAGGTGAGCGGGGACTCTCGG 895
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
284 ProValTyrValAsnGlnValLysValArgValSer-----AspAlaLeu 298
Qy 896 CGGAGCAGAG-----CAGAGCGCGAGCGCATGATGCTGGGATTT 937
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
299 LeuGluLysLysLysLeuMetAlaAlaArgGluArgLysAlaThrLysThrLeuGlyIle 318
Qy 938 CTGATTGGCGGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 997
Db ::::: ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
319 IleLeuGlyValPheIleValCysTrpLeuProPhePheIleIleSerLeuValMetPro 338
Qy 998 CTCTGT-----GCTGCGAGCTGCGCCCGCCATCTGGAAGACATATTTCTGTGCTGGC 1051
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
339 IleCysLysAspAlaCysTrpPheHisGlnAlaIlePheAspPhePheThrTrpLeuGly 358
Qy 1052 TACTCCATTTCTTCTTCAACCCCTGATTTACACAGCTTTTACACAGAACTACACAACT 1111
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
359 TyrValAsnSerLeuIleAsnProIleIleTyrThrMetSerAsnGluAspPheLysGln 378
Qy 1112 GCCTTCAAGAGCTC 1126
Db ||||| |||||
379 AlaPheHisLysLeu 383

RESULT 11
SHT1 DROME
ID -SHT1 DROME STANDARD; PRT; 564 AA.
AC P20905; OSV21;
DT 01-FEB-1991 (Rel. 17, Created)
DT 01-FEB-1991 (Rel. 17, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE 5-hydroxytryptamine receptor 1 (5-HT receptor) (Serotonin receptor).
GN 5-HT7 OR 5HT-R1 OR CG12073.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Oregon-R; TISSUE=Head;
RX MEDLINE=91062395; PubMed=2174167;
KW Witz P., Amalaki N., Plassat J.-L., Maroteaux L., Borrelli E., Hen R.;
RT "Cloning and characterization of a Drosophila serotonin receptor that
activates adenylylate cyclase.";

```

```

RL Proc. Natl. Acad. Sci. U.S.A. 87:8940-8944 (1990).
[2]
RN SEQUENCE FROM N.A.
RP STRAIN=Berkley;
RC MEDLINE=20196006; PubMed=10731132;
RX Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA Abrial J.F., Agbayani A., An H.-J., Andrews-Pfankoch C., Baldwin D.,
RA Balles R.M., Basay A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.V., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brottier P.,
RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
RA Foster C., Gabriellian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D.A., Heiman T.J., Hernandez J.R., Houck J.,
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Laoko P., Lei Y., Levitsky A.A., Li J.H., Li Z., Liang Y., Lin X.,
RA Liu X., Matti B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacle J.M.,
RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.-Y., Wassarman D.A., Weinstock G.M., Weissbach J.,
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT "The genome sequence of Drosophila melanogaster.";
RL Science 287:2185-2195 (2000).
CC -!- FUNCTION: This is one of the several different receptors for 5-
CC hydroxytryptamine (serotonin), a biogenic hormone that functions
CC as a neurotransmitter, a hormone, and a mitogen. The activity of
CC this receptor is mediated by G proteins which activate adenylylate
CC cyclase.
CC -!- SUBCELLULAR LOCATION: Integral membrane protein.
CC -!- TISSUE SPECIFICITY: Head.
CC -!- SIMILARITY: Belongs to family 1 of G-protein coupled receptors.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; M55533; AAA28305.1; -.
DR EMBL; AE003776; AAF57104.1; -.
DR FlyBase; FBgn0004573; 5-HT7
DR InterPro; IPR000276; GPCR_Rhodopsn.
DR InterPro; IPR007455; Serglycin.
DR Pfam; PF00001; 7tm1; 1.
DR Pfam; PF04360; Serglycin; 1.
DR PRINTS; PR00237; GPCR_RHODOPSN
DR PROSITE; PS00237; G_PROTEIN_RECEP_F1_1; 1.
DR PROSITE; PS00262; G_PROTEIN_RECEP_F2_1; 1.
KW G-protein coupled receptor; Transmembrane; Glycoprotein; Repeat.
FT TRANSMEM 29 51 0 (POTENTIAL).
FT TRANSMEM 165 188 1 (POTENTIAL).

```

FT DOMAIN 189 198 CYTOPLASMIC (POTENTIAL).
 FT TRANSMEM 199 222 2 (POTENTIAL).
 FT DOMAIN 223 236 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 237 258 3 (POTENTIAL).
 FT DOMAIN 259 278 CYTOPLASMIC (POTENTIAL).
 FT TRANSMEM 279 302 4 (POTENTIAL).
 FT DOMAIN 303 330 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 331 353 5 (POTENTIAL).
 FT DOMAIN 354 454 CYTOPLASMIC (POTENTIAL).
 FT TRANSMEM 455 476 6 (POTENTIAL).
 FT DOMAIN 477 487 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 488 510 7 (POTENTIAL).
 FT DOMAIN 511 564 CYTOPLASMIC (POTENTIAL).
 FT TRANSMEM 565 608 9 X 2 AA TANDEM REPEATS OF G-S.
 FT REPEAT 609 610 1.
 FT REPEAT 611 612 2.
 FT REPEAT 613 614 3.
 FT REPEAT 615 616 4.
 FT REPEAT 617 618 5.
 FT REPEAT 619 620 6.
 FT REPEAT 621 622 7.
 FT REPEAT 623 624 8.
 FT REPEAT 625 626 9.
 FT DISULFID 627 628 BY SIMILARITY.
 SQ SEQUENCE 564 AA; 60861 MW; 0C8B9F8DA63D8095 CRC64;

Alignment Scores:

Pred. No.: 9,12e-16 Length: 564
 Score: 441.50 Matches: 137
 Percent Similarity: 44.40% Conservative: 69
 Best Local Similarity: 29.53% Mismatches: 150
 Query Match: 20.50% Indels: 108
 DB: 1 Gaps: 12

US-09-976-782-15 (1-1152) x SHT1_DROME (1-564)

QY 11 GCGCTAGCCTTTCAGTGGCCACCGCGCGGTGGCTTCCCTGGGACCGGAGACCAGC 70
 Db SerAlaPheLeuGlyAlaIleAlaSerAlaSerThrGlySerGlySerGlySerGly 95
 QY 71 AGCGACCGCGGACCCACGAGCGGAGGAGGATCTCGGT-----TCGACC 115
 Db SerGlySerGlySerGlySerGlySerGlySerGlySerGlySerGlySerGlySer 115
 QY 116 CCGAGCGCGCGCTCTCGCGCGCGCGAGG----- 145
 Db 116 ProIle--AlaIleValSerTyrGlnGlyIleThrSerSerAsnLeuGlyAspSerAsn 134
 QY 145 ----- 145
 Db ThrThrLeuValProLeuSerAspThrProLeuLeuGluPheAlaAlaGlyGlu 154
 QY 146 -----CCGCCCTTCTCTCTTCACGGTCTGGTGGTGGTGGTGGTGGTGGTGGTGG 196
 Db PheValLeuProLeuThrSerIlePheValSerIleValLeuLeuIleValIleLeu 174
 QY 197 ATCGCTGCCACTTCTCGTGGACCGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGC 256
 Db 175 -----GlyThrValValGlyAsnValLeuValCysIleAlaValCysMetValArgLys 192
 QY 257 TTCACCGCGCGCGGATACCTGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGCA 316
 Db 193 LeuArgArgProCysAsnTyrLeuLeuValSerLeuLeuLeuSerAspLeuValAla 212
 QY 317 GCGCTGGCGATCCACCGAGCTGGCGAGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 376
 Db 213 LeuLeuValMetProMetAlaLeuLeuTyrGluValLeu-----GluLysTrpAsnPhe 230
 QY 377 GGCGGAGCGCTGTGGCACGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 436
 Db 231 GlyProLeuLeuCysAspIleTrpValSerPheAspVal-----LeuCysCysThr 247
 QY 437 GCGCGCTCGGGAACGTGGGGGCCATCGCGCTGGCGCGCGGCGGCGGCGGCGGCGGCG 496

Db 248 AlaSerIleLeuAsnLeuCysAlaIleSerValAspArgTyrLeuAlaIleThrLysPro 267
 QY 497 CTGCAGCACACGCTGGCGACCGCGCGCTCGTTCCTCATGATCGCGCTCGCGCGCG 556
 Db 268 LeuGluTyrGlyValLysArgThrProArgMetMetLeuValGlyIleValTrp 287
 QY 557 GTGCGCTGGCGCTCATCCCTCGCGCGCTCTTTGGCGGGCGGAGGTGTCGAC 616
 Db 288 LeuAlaAlaCysIleSerLeuProLeuLeuIleLeuGlyAsnGluHisGluAsp 307
 QY 617 GTCGCTCCAGCGC---TCGAGGTGACCGGGAACCTCTCTATCCGCTTCTCCACC 673
 Db 308 GluGluGlyGlnProIleCysThrValCysGlnAsnPheAlaTyrGlnIleTyrAlaThr 327
 QY 674 CGCGGCGCTTCCACCTGCGCTGGCGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 733
 Db 328 LeuGlySerPheTyrIleProLeuSerValMetLeuPheValTyrTyrGlnIlePheArg 347
 QY 734 GCGGCCAAGTTTCGTTTCGCGCGCGCGAGAGCTGTCTG----- 775
 Db 348 AlaAla-----ArgArgIleValLeuGluLysArgAlaGln 360
 QY 776 -----CCG 778
 Db 361 ThrHisLeuGlnAlaLeuAsnGlyThrGlySerProSerAlaProGlnAlaProPro 380
 QY 779 TTGCGCGCCACCATGCAGGTGAGTCCCAAGTAAAGGA----- 816
 Db 381 Leu--GlyHisThrGluLeuAlaSerSerGlyAsnGlyGlnArgHisSerValGlyA 400
 QY 817 -----AGCACCTGATGAGGTGAAAGTGGTGTTCACGCGCACATTGCAG-- 858
 Db 400 snThrSerLeuThrTyrSerThrCysGlyLeuSerSerGlyGlyGlyAlaLeuAlaG 420
 QY 859 -----AGCAACGCTGTCTTCAGGTGAGCGGGACTCTCTGGCGGGA----- 900
 Db 420 LysHisGlySerGlyGlyGlyValSerGlySerThrGlyLeuLeuGlySerProHisL 440
 QY 901 -----GCAGAGGAGGCGGAGCGGAGCCATGATGTGGGAATTCG 940
 Db 440 ysLysLeuArgPheGlnLeuAla--LysGluLysLysAlaSerThrThrLeuGlyIle 459
 QY 941 ATTGGCGT 1000
 Db 460 MetSerAlaPheThrValCysTrpLeuProPhePheIleLeuAlaLeuIleArgProPhe 479
 QY 1001 TGTGCTGCGAGCTGCCCCCATCTGGAAAGCATATTTCTGTGGTGGTGGTGGTGGTGG 1060
 Db 480 GluThrMetHisValProAlaSerLeuSerSerLeuPheLeuTrpLeuGlyTyrAlaAsn 499
 QY 1061 TCTTTTCTTCAACCCCTGATTTACACAGCTTTTAAACAAGACTACACAATGCCTTCAAG 1120
 Db 500 SerLeuLeuAsnProIleIleTyrAlaThrLeuAsnArgAspPheArgLysProPheGln 519
 QY 1121 AGCCTC 1126
 Db 520 GluIle 521
 RESULT 12
 SHID_CANFA
 ID SHID CANFA STANDARD; PRT; 377 AA.
 AC P11614;
 DT 01-OCT-1989 (Rel. 12, Created)
 DT 01-OCT-1989 (Rel. 12, Last sequence update)
 DT 10-OCT-2003 (Rel. 42, Last annotation update)
 DE 5-hydroxytryptamine 1D receptor (5-HT-1D) (Serotonin receptor).
 GN HTR1D OR RDC4.
 OS Canis familiaris (Dog).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.
 NCBI_TaxID=9615;
 RN [1]

SEQUENCE FROM N.A.
 RP TISSUE=Thyroid;
 RC MEDLINE=89242119; PubMed=2541503;
 RX Libert F., Parmentier M., Lefort A., Dinsart C., van Sande J.,
 RA Maenhaut C., Simons M.-J., Dumont J.E., Vassart G.;
 RT "Selective amplification and cloning of four new members of the G
 RL protein-coupled receptor family";
 RN Science 244:569-572(1989).
 [2]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Thyroid;
 RX MEDLINE=90245610; PubMed=2159630;
 RA Libert F., Parmentier M., Lefort A., Dumont J.E., Vassart G.;
 RT "Complete nucleotide sequence of a putative G protein coupled
 RL receptor: RDC4";
 RN Nucleic Acids Res. 18:1916-1916(1990).
 [3]
 RP MEDLINE=92062181; PubMed=1659418;
 RX Maenhaut C., van Sande J., Massart C., Dinsart C., Libert F.,
 RA Monferini E., Giraldo E., Ladinsky H., Vassart G., Dumont J.E.;
 RT "The orphan receptor cDNA RDC4 encodes a 5-HT1D serotonin receptor";
 RL Biochem. Biophys. Res. Commun. 180:1460-1468(1991).
 [4]
 RP MEDLINE=92100052; PubMed=1758439;
 RX Zgombick J.M., Weinshank R.L., Macchi M., Schechter L.E.,
 RA Branchek T.A., Hartig P.R.;
 RT "Expression and pharmacological characterization of a canine 5-
 RL hydroxytryptamine receptor subtype";
 RN Mol. Pharmacol. 40:1036-1042(1991).
 CC -!- FUNCTION: This is one of the several different receptors for 5-
 CC hydroxytryptamine (serotonin), a biogenic hormone that functions
 CC as a neurotransmitter, a hormone, and a mitogen. The activity of
 CC this receptor is mediated by G proteins that inhibit adenylyate
 CC cyclase activity.
 CC -!- SUBCELLULAR LOCATION: Integral membrane protein.
 CC -!- SIMILARITY: Belongs to family 1 of G-protein coupled receptors.
 CC Strongest to the other 5HT-1 subtype receptors.

 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

 CC EMBL; X14049; CAA32207.1; -;
 DR PIR; B30341; B30341.
 DR InterPro; IPR000276; GPCR_Rhodpsn.
 DR Pfam; PF00001; 7tm1; 1.
 DR PRINTS; PR00237; GPCRHHODPSN.
 DR PROSITE; PS00237; G-PROTEIN RECF1_1; 1.
 DR PROSITE; PS00262; G-PROTEIN RECF1_2; 1.
 KW G-protein coupled receptor; Transmembrane; Glycoprotein;
 KW Multigene family.
 KW DOMAIN 1 38 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 39 62 1 (POTENTIAL).
 FT DOMAIN 63 75 CYTOPLASMIC (POTENTIAL).
 FT TRANSMEM 76 98 2 (POTENTIAL).
 FT DOMAIN 99 108 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 109 134 3 (POTENTIAL).
 FT DOMAIN 135 154 CYTOPLASMIC (POTENTIAL).
 FT TRANSMEM 155 176 4 (POTENTIAL).
 FT DOMAIN 177 194 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 195 218 5 (POTENTIAL).
 FT DOMAIN 219 302 CYTOPLASMIC (POTENTIAL).
 FT TRANSMEM 303 326 6 (POTENTIAL).
 FT DOMAIN 327 335 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 336 360 7 (POTENTIAL).
 FT DOMAIN 361 377 CYTOPLASMIC (POTENTIAL).
 FT CARBOHYD 5 N-LINKED (GLCNAC...) (POTENTIAL).
 FT CARBOHYD 17 N-LINKED (GLCNAC...) (POTENTIAL).
 FT CARBOHYD 21 N-LINKED (GLCNAC...) (POTENTIAL).
 FT DISULFID 111 188 BY SIMILARITY.
 SQ SEQUENCE 377 AA; 41882 MW; 856406DCB2123EE4 CRC64;
 Alignment Scores:
 Pred. No.: 9,78e-16 Length: 377
 Score: 441.00 Matches: 121
 Percent Similarity: 47.10% Conservative: 66
 Best Local Similarity: 30.48% Mismatches: 142
 Query Match: 20.47% Indels: 68
 DB: 1 Gaps: 11
 US-09-976-782-15 (1-1152) x 5H1D_CANFA (1-377)
 QY 8 GAGCGCGCTAGCCTTTTCAGTGGCCACCGCGCGGTTCCTTGCCTTGCCTGGGACCCGAGACC 67
 DB |||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||: 33
 14 GluAlaSerAsnArgSerLeuAsnAlaThrGluThrProGluAlaTrpGlyProGluThr 33
 QY 68 AGCAGCGACCCGGGACCCCAAGCCCGAGAGGATACCTCGGTTCCGACCCGAGCGGCC 127
 DB ----- 33
 33 ----- 33
 QY 128 GTCCTGCGGCGCGAGGCGCCCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 187
 DB -----LeuGlnAlaLeuLysIleSerLeuAlaLeuLeu 45
 34 -----LeuGlnAlaLeuLysIleSerLeuAlaLeuLeu 45
 QY 188 GTGCTGTGATCGCTGCGACCTTCTGTGGAACTGCTGGTTCGGTACCATCCCGCGG 247
 DB :|||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||: 65
 46 SerIleIleThrMetAlaThrAlaLeuSerAsnAlaPheValLeuThrIlePheLeu 65
 QY 248 GTCGTGCTTCCACCGCGTCCGCAATACTTGTGGCTCGACGGCGCTCGGACGAA 307
 DB |||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||: 85
 66 ThrArgLysLeuHisThrProAlaAsnTyrLeuIleGlySerLeuAlaMetThrAspLeu 85
 QY 308 CTAGTGGCAGCGTGGCGATCCACCGAGCTGGCGAGTGTGTCTCGACCGCGGCGGT 367
 DB |||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||: 103
 86 LeuValSerIleLeuValMetProIleSerIleAla-----TyrThrThrArgThr 103
 QY 368 CGGCTGTGGCGCGAGCGCTGTGCACAGTGTGGATCTCTCTCGACCGCGGAGCTGTCTG 427
 DB |||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||: 120
 104 TrpSerPheGlyGlnIleLeuCysAspIleTrpLeuSerSerAsp-----IleThr 120
 QY 428 TGCTGCCCCCGCGCTCGGGAAGTGGCGCCATCGCCCTGGCGCGACGGGCGCCATC 487
 DB |||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||: 140
 121 CysCysThrAlaSerIleLeuHisLeuCysValIleAlaLeuAspArgTyrTrpAlaIle 140
 QY 488 ACAGCGCACCTCGACACAGCTGCGCACCGCGCGCGCTCTCTCTCATATGATCGCG 547
 DB |||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||: 160
 141 ThrAspAlaLeuGluTyrSerLysArgArgThrAlaGlyArgAlaAlaValMetIleAla 160
 QY 548 CTCGCCCGGTGCGCGCTGTCCCTCTCGCGCGCGCTCTCTCTCTCTCTCTCTCTCTCTCT 607
 DB |||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||: 179
 161 ThrValTrpValIleSerIleCysIleSerIleProPro-----LeuPheTrpArgGlnAla 179
 QY 608 GTGTGCGACGCTCGGCTCCAGCGTCCAGCTG---AGCCGGGACCCCTCTATCCGCGC 664
 DB |||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||: 199
 180 LysAlaGlnGluAspMetSerAspCysGlnValAsnThrSerGlnIleSerTyrThrIle 199
 QY 665 TTCTCCACCGCGCGCTTCCACTGCGCTGTGGCGTGGCGTGGTGGTGGTGGTGGTGGTGGT 724
 DB :|||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||: 219
 200 TyrSerThrCysGlyAlaPheTyrIleProSerValLeuLeuIleLeuTyrGlyArg 219
 QY 725 ATCTACGAGGCGCGCAAGTTTCGT-----TTGCGCGCGCGCGCGG 763
 DB |||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||: 239
 220 IleTyrValAlaAlaArgAsnArgIleLeuAsnProProSerLeuTyrGlyLysArgPhe 239
 QY 764 AGAGCTGTGCTGCCGTG-----CCGCGC 787
 DB :|||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||: 259
 240 ThrThrAlaGlnLeuIleThrGlySerAlaGlySerSerLeuSerLeuSerProSer 259
 QY 788 ACCATGACAGTGAGGTCCCAAGGTAAAGGACCATGATGAGGCTGAAGTGTGTACG 847
 DB :|||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||: 847


```
QY 620 CGGCTCCAGCGTGCCAGGTGAGCGGAA---CCCTCTATGCGCGCTTCTCCACCGCG 676
Db   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 677 GCGCGCTCCAGCTCGCGCTGGCGTGCGGTGCTACCGGAAGATCAGCGCG 736
Db   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 737 GCCAGTTTCGTTTC-----GCGCGCGC---CGGAGACT 769
Db   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 812 AAGNAGCACCTGATGAGCT---GAAGTGTGTTTCAGCGGCACATGCAAGCAACGGTG 868
Db   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 869 TCCTTCCAGGTGAGCGGAGCTCCTCGCGGAGCAGAG-----GAG 910
Db   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 911 AGCGAGCAGCAGCATGATGCGGAATCTGATGCGCGTGTGCTGTGTCGATCCCG 970
Db   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 971 TTCCTTCGAGGAACCTATCAGCCACTCTGT-----GCTGAGCGCTGCCCGCCATC 1024
Db   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1025 TGAAGAACATATTTCTGCGTTCGCTGCTACTCTCAATCTTCTTCAACCCCTGATTTAC 1084
Db   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1085 ACAGCTTTTAAAGAACTACAAATGCTTCAAGAGCCTC 1126
Db   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 366 ThrMetSerAsnGluAspPheLysGlnAlaPheHisLysLeu 379
Db   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

RESULT 15
SH1D_CAVPO STANDARD; PRT; 376 AA.
AC Q60484; O08891;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE 5-Hydroxytryptamine 1D receptor (5-HT-1D) (Serotonin receptor).
GN HTR1D.
OS Cavia porcellus (Guinea pig).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Hystriognathi; Caviidae; Cavia.
OX NCBI_TaxID=10141;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Hartley; TISSUE=Brain;
RX MEDLINE=97133356; PubMed=8978753;
RA Wurch T., Palmier C., Colpaert F.C., Pauwels P.J.;
RT "Sequence and functional analysis of cloned guinea pig and rat
RT serotonin 5-HT1D receptors: common pharmacological features within
RT the 5-HT1D receptor subfamily.";
RL J. Neurochem. 68:410-418(1997).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Liver;
RX MEDLINE=97368662; PubMed=9225276;
RA Zgombick J.M., Bard J.A., Kucharewicz S.A., Urquhart D.A.,
RA Weinschank R.L., Branchek T.A.;
RT "Molecular cloning and pharmacological characterization of guinea pig
RT 5-HT1B and 5-HT1D receptors.";
RL Neuropharmacology 36:513-524(1997).
CC -!- FUNCTION: This is one of the several different receptors for 5-
CC hydroxytryptamine (serotonin), a biogenic hormone that functions
```

```
CC as a neurotransmitter, a hormone, and a mitogen. The activity of
CC this receptor is mediated by G proteins that inhibit adenylate
CC cyclase activity.
CC -!- SUBCELLULAR LOCATION: Integral membrane protein.
CC -!- SIMILARITY: Belongs to family 1 of G-protein coupled receptors.
CC Strongest to the other 5HT-1 subtype receptors.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch)
CC
CC EMBL; X94436; CA64210.1; -.
CC EMBL; U82174; AAB58499.1; -.
CC InterPro; IPR000276; GPCR_Rhodpsn.
CC Pfam; PF00001; 7tm.1; 1.
CC PRINTS; PR00237; GPCRHHODOPS.
CC PROSITE; PS00237; G_PROTEIN_RECP_F1_1; 1.
CC PROSITE; PS00262; G_PROTEIN_RECP_F1_2; 1.
CC G-protein coupled receptor; Transmembrane; Glycoprotein;
CC Multigene family.
CC
CC FT DOMAIN 1 38 EXTRACELLULAR (POTENTIAL).
CC FT TRANSMEM 39 62 1 (POTENTIAL).
CC FT DOMAIN 63 75 CYTOPLASMIC (POTENTIAL).
CC FT TRANSMEM 76 98 2 (POTENTIAL).
CC FT DOMAIN 99 108 EXTRACELLULAR (POTENTIAL).
CC FT TRANSMEM 109 134 3 (POTENTIAL).
CC FT DOMAIN 135 154 CYTOPLASMIC (POTENTIAL).
CC FT TRANSMEM 155 176 4 (POTENTIAL).
CC FT DOMAIN 177 194 EXTRACELLULAR (POTENTIAL).
CC FT TRANSMEM 195 218 5 (POTENTIAL).
CC FT DOMAIN 219 301 CYTOPLASMIC (POTENTIAL).
CC FT TRANSMEM 302 324 6 (POTENTIAL).
CC FT DOMAIN 325 334 EXTRACELLULAR (POTENTIAL).
CC FT TRANSMEM 335 359 7 (POTENTIAL).
CC FT DOMAIN 360 376 CYTOPLASMIC (POTENTIAL).
CC FT CARBOHYD 5 5 N-LINKED (GLCNAC. .) (POTENTIAL).
CC FT CARBOHYD 17 17 N-LINKED (GLCNAC. .) (POTENTIAL).
CC FT DISULFID 111 188 N-LINKED (GLCNAC. .) (POTENTIAL).
CC FT CONFLICT 222 222 BY SIMILARITY.
CC FT CONFLICT 222 222 R -> A (IN REF. 2).
CC SQ SEQUENCE 376 AA; 41768 MW; 531D4E1C2C819035 CRC64;
CC
CC Alignment Scores:
CC Pred. No.: 1.47e-15 Length: 376
CC Score: 437.50 Matches: 124
CC Percent Similarity: 45.98% Conservative: 59
CC Best Local Similarity: 31.16% Mismatches: 146
CC Query Match: 20.31% Indels: 69
CC DB: 1 Gaps: 12
CC
CC US-09-976-782-15 (1-1152) x SH1D_CAVPO (1-376)
CC
CC QY 53 CTGGAGCCGAGCAGCAGCAGCGCGCGCCCA-----AGCCGAGAGGATA 103
CC   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
CC Db 1 MetSerProAsnGlnSerGluGluLeuProGlnGluAlaSerAsnArgSerLeu 20
CC
CC QY 104 CTCGGTTTCGACCGCGCGCGCTCTCGCGCGCGCGCGCGCGCTCTCTGTCTTC 163
CC   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
CC Db 21 AsnAlaThrGluThrProGlyAspTrpAspProGlyLeuLeuGlnAlaLeuIysValSer 40
CC
CC QY 164 ACGTCTCTGGTGTGACGCTCTAGTCTGTGTGATCGCTGCACCTTCCTGTGGAACCTG 223
CC   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
CC Db 41 LeuValValValLeuSerIleIleThrLeu-----AlaThrValLeuSerAsnAla 57
CC
CC QY 224 CTGTTTCGGTTCACCATCCCGCGGTTCCTCCACCGCGTCCGCATTAACCTGTGTG 283
CC   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
CC Db 58 PheValLeuThrThrIleLeuLeuLeuThrArgLysLeuHisThrProAlaAsnTyrLeu 77
CC
CC 284 GCCTCGAGCGCGCTCTCGGACGAACTAGTGAGCGCTGGCGATGCCACCGAGCCTGGCG 343
```


Db	78	GlySerLeuAlaThrThrAspLeuLeuValSerIleLeuValMetProIleSerIleAla	97
Qy	344	AGTGAGCTGTCCACGGCGGACAGTCGGCTGTGGCGCGGACCGCTGTGCACGTGTGGATC	403
Db	98	-----TyrThrThrThrArgThrTriaAsnPhgIleGlyGlnIleLeuCysAspIleTirpVal	115
Qy	404	TCCTTCAGCGCGGAGGCTGTCTGTGTGTGCCCCCGCGCCCTCGCGGAACGTGGCGCCATC	463
Db	116	SerSerAsp-----IleThrCysCysThrAlaSerIleLeuHisLeuCysValIle	132
Qy	464	GCCTGGCGCGACGGGGCCATCACACGGCACCTGCAGCACACGCTCGCACCCGGCAGC	523
Db	133	AlaLeuAspArgTyrTirpAlaIleThrAspAlaLeuGluIyrSerIysArgArgThrAla	152
Qy	524	CGCGCCTGTGTCTATGATGGCGCTCGCCGGGTGCCGTGGCGCTCATGCCCTCGCG	583
Db	153	GlyHisAlaGlyAlaMetIleAlaAlaValTirpValIleSerIleCysIleSerIlePro	172
Qy	584	CGCGTGTCTTTCGGCGGGCGAGGTGGCGAGCTCGGCTCCAGCGCTGCCAGGTG---	640
Db	173	Pro---LeuPheTirpArgIleAlaGlnAlaGlnGluMetSerAspCysLeuValAsn	191
Qy	641	AGCGGGAACCTCCTATGCGCCTCTCCACCCGCGCGCTTCCACCTGCCCTTGGC	700
Db	192	ThrSerGlnIleSerTyrThrIleTyrSerThrCysGlyAlaPheTyrIleProSerVal	211
Qy	701	GTGGTGGCGTTTGTCTACGGAAGATCTACAGCGGGCCAAAGTTTCGTTTCGGCGCGCGC	760
Db	212	LeuLeuIleIleLeuTyrSerArgIleTyrArgAlaAla-----	224
Qy	761	CGGAGAGCTGTGTCGCGTTGCCGGCCACCATGCAGGTGAGGTCCCAAGTAAAGGAACGA	820
Db	225	-----ArgSerArgIleLeuAsnPro	231
Qy	821	CCTGATGAGCTGAAGTGGTGTTC---ACGCGCAT	853
Db	232	ProSerLeuSerGlyLysArgPheThrThrAlaHisLeuIleThrGlySerAlaGlySer	251
Qy	854	-----TGCAAGCAACGGTGTCTTCCAGGTGAC-	883
Db	252	SerLeuCysSerLeuAsnProSerLeuHisGluGlyHisMethIisProGlySerProLeu	271
Qy	884	-----GGGACCTCTCGGGGAGCAGAAG-----	907
Db	272	PhePheAsnHisValArgIleLysLeuAlaAspSerValIleGluArgLysArgIleSer	291
Qy	908	-----GAGAGCGGAGCAGCCATGATGGTGGGAATTCGATTCGCGTGTGTGTGCTG	958
Db	292	AlaAlaArgGluArgLysAlaThrLysThrLeuGlyIleIleLeuGlyAlaPheIleVal	311
Qy	959	TGCTGGATCCCTTCTTCGACGGAACTCATAGCCCACTCTGT-----GCTCGAGC	1012
Db	312	CysTirpLeuProPhePheValSerLeuValLeuProIleCysArgAspSerCysTirp	331
Qy	1013	CTGCCCCCATCTGGAAAAACATATTCCTGTGGCTTGCTACTCCCAATCTTTTCTCAAC	1072
Db	332	IleHisProAlaLeuPheAspPheThrThrTirpLeuGlyTyrLeuAsnSerLeuIleAsn	351
Qy	1073	CCCTGATTACACAGCTTTTAAACAGAACTACAACAATGCCTTCAGAGCCTC	1126
Db	352	ProIleIleTyrThrValPheAsnGluAspPheArgGlnAlaPheGlnLysVal	369

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - protein search, using frame_plus_n2p model

Run on: August 31, 2004, 19:58:47 ; Search time 148.5 Seconds
(without alignments)
4383.767 Million cell updates/sec

Title: US-09-976-782-15

Perfect score: 2154

Sequence: 1 cgcctggaggccgctagcc.....aagcagatgaacacagg 1152

Scoring table: BLOSUM62

Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 1586107 seqs, 282547505 residues

Total number of hits satisfying chosen parameters: 3172214

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+ n2p.model -DEV=xlp
-Q=/cgn2_1/USPRO.spool.P/US09976782/runat 31082004_115949_10296/app.query.fasta_1.1351
-DB=A_Geneseq 29Jan04 -QFMT=fastan -SUFFIX=n2p.rag -MINMATCH=0.1 -LOOPCL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=blos62 -TRANS=human40.cdi
-LIST=45 -DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USERS=US09976782 @CGN 1_1_224 @runat 31082004_115949_10296 -NCPU=6 -ICPU=3
-NO MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FCGAPOP=6
-FCGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : A_Geneseq 29Jan04.*

1: Geneseqp1980s.*

2: Geneseqp1990s.*

3: Geneseqp2000s.*

4: Geneseqp2001s.*

5: Geneseqp2002s.*

6: Geneseqp2003as.*

7: Geneseqp2003bs.*

8: Geneseqp2004s.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1946	90.3	379	5 ABB78809	Abb78809 Human NOV
2	1861.5	86.4	370	5 AAE18654	Aae18654 Human G-p
3	1860.5	86.4	370	4 AAM47211	Aam47211 Human NOV
4	1857.5	86.2	370	4 AAM47212	Aam47212 Human NOV
5	1856.5	86.2	372	5 AAE15638	Aae15638 Human G-p
6	1819.5	84.5	380	5 ABG60235	Abg60235 Human hyd
7	1453.5	67.5	370	2 AAR58686	Aar58686 Rat MR22
8	1453.5	67.5	370	7 ADE56994	Ade56994 Rat Prote
9	1452.5	67.4	370	2 AAR57066	Aar57066 Murine se
10	1452.5	67.4	370	6 AAE31673	Aae31673 Mouse 5-H

11	1296.5	60.2	606	7 ADC86157	Adc86157 Human GPC
12	1063.5	49.4	357	5 AAU79252	Aau79252 Human 5-h
13	1063.5	49.4	357	6 ABP81980	Abp81980 Human 5-H
14	1063.5	49.4	357	7 AAE38595	Aae38595 Human 5-H
15	1063.5	49.4	357	7 ADE56996	Ade56996 Human Pro
16	1058.5	49.1	357	4 ABB56327	Abb56327 Non-endog
17	1053.5	48.9	357	5 AAU79253	Aau79253 Human 5-h
18	1022.5	47.5	357	5 AAR45848	Aar45848 Human 5HT
19	1022.5	47.5	357	5 AAU75167	Aau75167 Mouse 5-H
20	1017.5	47.1	357	2 AAR45847	Aar45847 Murine 5H
21	1014.5	47.1	357	2 AAR58685	Aar58685 Rat REC17
22	647.5	30.1	212	4 AAG62846	Aag62846 Amino aci
23	596	27.7	111	2 AAR57067	Aar57067 Human ser
24	501	23.3	99	4 AAG62845	Aag62845 Amino aci
25	448	20.8	390	2 AAY28304	Aay28304 Serotonin
26	448	20.8	390	4 AAE00911	Aae00911 Human 5-h
27	448	20.8	390	6 ABG73850	Abg73850 Human ser
28	448	20.8	390	6 ABP81761	Abp81761 Human 5-H
29	448	20.8	390	7 ADE40461	Ade40461 Human ser
30	447	20.8	390	2 AAR43060	Aar43060 Human 5-H
31	447	20.8	390	4 ABB56317	Abb56317 Non-endog
32	444	20.6	90	6 ABP72590	Abp72590 Human LP3
33	441.5	20.5	564	4 ABB58411	Abb58411 Drosophil
34	441.5	20.5	564	7 AAE38198	Aae38198 Fruit fly
35	441	20.5	377	6 ABG74161	Abg74161 Dog 5-Ht
36	438	20.3	386	5 ABB57292	Abb57292 Mouse isc
37	431.5	20.0	376	2 AAR15137	Aar15137 Human ser
38	430	20.0	377	2 AAY28303	Aay28303 Serotonin
39	430	20.0	377	4 AAB47185	Aab47185 5-hydroxy
40	430	20.0	377	6 ABG73849	Abg73849 Human ser
41	430	20.0	377	6 ABP81762	Abp81762 Human 5-H
42	430	20.0	377	7 ADE58162	Ade58162 Human pro
43	429	19.9	377	4 ABB56318	Abb56318 Non-endog
44	428	19.9	422	4 AAB70249	Aab70249 HTR1A pro
45	428	19.9	451	5 ABB75677	Abg75677 Human 5HT

ALIGNMENTS

RESULT 1

ABB78809

ID ABB78809 standard; protein; 379 AA.

XX ABB78809;

XX AC

DT 29-JUL-2002 (first entry)

XX

DE Human NOV5 protein sequence SEQ ID NO:16.

XX

Human; NOV5; cytostatic; antiarteriosclerotic; cardiovascular; lymphoma; anti-diabetic; immunosuppressive; neuroprotective; gene therapy; cancer; cardiomyopathy; atherosclerosis; cell signal processing; diabetes; AIDS; metabolic pathway modulation; neoplastic; neurological disorder; asthma; adenocarcinoma; prostate cancer; uterus cancer; immune response; Crohn's disease; multiple sclerosis; Graft versus host disease; chromosome 2.

XX Homo sapiens.

XX W0200230974-A2.

XX 18-APR-2002.

XX 12-OCT-2001; 2001WO-US031922.

XX 12-OCT-2000; 2000US-0240113P.

XX 16-OCT-2000; 2000US-0240625P.

XX 16-OCT-2000; 2000US-0240637P.

XX 16-OCT-2000; 2000US-0240648P.

XX 16-OCT-2000; 2000US-0240662P.

XX 16-OCT-2000; 2000US-0240669P.

XX 16-OCT-2000; 2000US-0240703P.

XX 16-OCT-2000; 2000US-0240732P.

QY 665 TTCTCCACCCGGCGGCTTCCACCTGCGCTTGGCGTGGTGGCGTGTGTCTACCGGAG 724
 Db 216 PheSerThrArgGlyAlaPheHisLeuProLeuGlyValValProPheValTyrArgLys 235
 QY 725 ATCTACGAGCGGCGCAAGCTTTCTGTTCCGCGCGCGGAGAGCTGTCTCGCGTTCGCG 784
 Db 236 IleTyrGluAlaAlaLysPheArgPheGlyArgArgArgAlaValLeuProLeuPro 255
 QY 785 GCCACCATGCGAGTGAGTCCAAAGTAAAGGAGACCTGATGAGCTGAAGTGGTTC 844
 Db 256 AlaThr-----SerLysValLysGluAlaProAspGluAlaGluValValPhe 271
 QY 845 ACGGCACATGTCAGGAGTCCCTTCCAGGTGAGCGGACTCTCTGCGGGAGCAG 904
 Db 272 ThrAlaHisCysLysAlaThrValSerPheGlnValSerGlyAspSerTrpArgGluGln 291
 QY 905 AAGGAGCGGAGCAGCATCATGTTGGAAATTCGATGTCGTTGTTGCTGTGCTGG 964
 Db 292 LysGluArgArgAlaAlaMetMetValGlyIleLeuIleGlyValPheValLeuCysTrp 311
 QY 965 ATCCCTCTCTCTGACGGAATCATCAGCCACTCTGTGCTGTGAGCTGCCCCCATC 1024
 Db 312 IleProPhePheLeuThrGluLeuIleSerProLeuCysAlaCysSerLeuProPhe 331
 QY 1025 TGGAAAGCATATTCTGTGCTGGCTTGGCTACTCCAATCTTTCTCAACCCCTGATTTC 1084
 Db 332 TrpLysSerIlePheLeuTrpLeuGlyTyrSerAsnSerPhePheAsnProLeuIleTyr 351
 QY 1085 ACAGCTTTTAAAGAACTCAACAATGCTTCAAGACCTCTTTACTTAAGCAGAGA 1141
 Db 352 ThrAlaPheAsnLysAsnTyrAsnAsnAlaPheLysSerLeuPheThrLysGlnArg 370

RESULT 4
 AAM47212
 ID AAM47212 standard; protein; 370 AA.
 XX AC AAM47212;
 AC AC
 DT 12-FEB-2002 (first entry)
 XX AC
 DE Human NOV5b protein.
 XX KW Human; NOV5b; metabolic disorder; neurodegenerative disorder;
 KW immune disorder; haematopoietic disorder; developmental disease; cancer;
 KW retinal disease; feeding disorder; vaccine; infection; gene therapy;
 KW neurological disorder; psychotic disorder; G-protein coupled receptor;
 KW analgesic; antidiabetic; viricide; neuroprotective; nootropic;
 KW anesthetic; antidepressant; antimigraine; anticonvulsant; neuroleptic;
 KW antipruritic; antiallergic; antiinflammatory; anorectic; antiarthritic;
 KW antipruritic; antithrombotic; antibacterial; fungicide;
 KW osteopathic; protozoacide; antiulcer; hypertensive; hypotensive;
 KW antiinfertility; vulnary; nephrotropic; antilipemic;
 KW serotonin receptor; chromosome 2.
 XX OS Homo sapiens.
 XX PH
 PH Key Location/Qualifiers
 FT Peptide 1..24
 FT /label= signal_peptide
 FT Protein 25..370
 FT /label= mature_NOV5b
 XX WO200174851-A2.
 XX 11-OCT-2001.
 XX 30-MAR-2001; 2001WO-US010039.
 XX 30-MAR-2000; 2000US-0193205P.
 XX 30-MAR-2000; 2000US-0193339P.
 XX 05-APR-2000; 2000US-0195343P.
 XX 06-APR-2000; 2000US-0195005P.

PR 06-APR-2000; 2000US-0195088P.
 PR 10-APR-2000; 2000US-0195792P.
 PR 11-APR-2000; 2000US-0196556P.
 PR 13-APR-2000; 2000US-0197081P.
 PR 14-APR-2000; 2000US-0197087P.
 PR 14-APR-2000; 2000US-0197525P.
 PR 29-MAR-2001; 2001US-00823187.
 XX (CURA-) CURAGEN CORP.
 PA Majumder K, Spaderna SK, Taupier RJ, Padigaru M, Burgess CE;
 PI Shinkets RA, Spytek KA, Liu X, Patturajan M, Gusev VV;
 XX WPI; 2001-626379/72.
 DR N-PSDB; ABA01985.
 XX
 PT New G protein-coupled receptor related polypeptides and polynucleotides
 PT for diagnosis, prevention and treatment of metabolic, neurodegenerative,
 PT retinal, immune, hematopoietic disorders, diabetes, obesity and
 PT infections.
 XX
 PS Claim 1; Page 47; 194pp; English.
 CC The present invention provides the protein and coding sequences of novel
 CC human G-protein coupled receptors, designated NOV1, NOV2, NOV3, NOV4,
 CC NOV5a, NOV5b, NOV6a, NOV6b, NOV7, NOV8a, NOV8b, NOV9 and NOV10. These can
 CC be used in the treatment of NOVX related diseases, including cancer,
 CC metabolic, neurodegenerative, immune, haematopoietic, developmental,
 CC retinal, feeding, neurological and psychotic diseases and disorders and
 CC infections. The present sequence is the NOV5b protein, the gene for which
 CC is found on chromosome 2. The NOV5b protein shares homology with
 CC serotonin receptors
 XX
 SQ Sequence 370 AA;

Alignment Scores:
 Pred. No.: 3,82e-136 Length: 370
 Score: 1857.50 Matches: 369
 Percent Similarity: 97.36% Conservative: 0
 Best Local Similarity: 97.36% Mismatches: 1
 Query Match: 86.23% Indels: 9
 DB: 4 Gaps: 3

US-09-976-782-15 (1-1152) x AAM47212 (1-370)

QY 5 ATGAGAGCGCTAGCCTTTAGTGGCCACCGCGCGCTGTCCTTGCCTGGGACCGAG 64
 Db 1 MetGluAlaAlaSerLeuSerValAlaThrAlaGlyValAlaLeuAlaLeuGlyProGlu 20
 QY 65 ACCAGCAGCGGACCGCGGACCCCAAGCCGAGAGGATACCTCGGTTGACCCCGAGCGGC 124
 Db 21 ThrSerSer-----GlyThrProSerProArgGlyIleLeuGlySerThrProSerGly 38
 QY 125 GCCGTCTGCGGCGCGAGGCGCCCTTCTCTCTTCTACGCTCTGCTGGTGGACGCTG 184
 Db 39 AlaValLeuProGlyArgGlyProProPheSerValPheThrValLeuValValThrLeu 58
 QY 185 CTAGTGTGTGATCGCTGCCACTTCTCTGTGGAACCTGCTGGTTCGGTCCGTCACATCCG 244
 Db 59 LeuValLeuLeuIleAlaAlaThrPheLeuTrpAsnLeuLeuValProValThrIlePro 78
 QY 245 CGGTCGCTGCTTCCACCGCGTCCGATTAACCTTGGTGGCTCGAGCGCGCTCGGAC 304
 Db 79 ArgValArgAlaPheHisArgValProHisAsnLeuValAlaSerThrAlaValSerAsp 98
 QY 305 GAAGTGTGCGGCTGCGGATGCCACCGAGCTGCGAGTGGAGTGTGAGTGTGAGCGGCGA 364
 Db 99 GluLeuValAlaAlaLeuAlaMetProProSerLeuAlaSerGluLeuSerThrGlyArg 118
 QY 365 CGTGTGCTGCGGCGGAGCCTGTGCGACGTGTGATCTCTTCGACGCGGAGCCTGT 424
 Db 119 ArgArgLeuLeuGlyArgSerLeuCysHisValTrpIleSerPheAspAla----- 135

QY 185 CTAGTGTGCTGATCGTGGCCACTTCTCTGTGGAACCTGCTGGTTCGGTCCACCATCCG 244
DB 59 LeuValLeuLeuLeuAlaAlaThrPheLeuTrpAsnLeuValProValThrIlePro 78
QY 245 CGGTCCTGCTCTCCACCGCGTGGCGGATCACTGTGGGCTCGACGGCGCTTCGAC 304
DB 79 ArgValArgAlaPheHisArgValProHisAsnLeuValAlaSerThrAlaValSerAsp 98
QY 305 GAACCTAGTGGCAGCGTGGCGATGCCACCGAGCTGGCGAGTGGAGTGTCTGACCGGGCGA 364
DB 99 GluLeuValAlaAlaLeuAlaMetProProSerLeuAlaSerGluLeuSerThrGlyArg 118
QY 365 CGTGGCTGTGTCGGCGGAGCCTGTGCCACCTGTGGATCTCTTCGACCGCGGAGCCTGT 424
DB 119 ArgArgLeuLeuGlyArgSerLeuCysHisValTrpIleSerPheHisVal----- 135
QY 425 CTGTGTCCTCCCGCGCTCGGGAACGTGGCGGCTATCGCCCTGGGCGCGAGCGGGGCC 484
DB 136 LeuCysCysProAlaGlyLeuGlyAsnValAlaAlaIleAlaLeuGlyArgAspGlyAla 155
QY 485 ATCACAGCGACCGTGGCGAGCACACGCTGGCGACCGCGCGCTCGTTCCTATGATC 544
DB 156 IleThrArgHisLeuGlnHisThrLeuArgThrArgSerArgAlaSerLeuLeuMetIle 175
QY 545 GCGCTCGCCCGGTCGCGTCCGCTCATCGCCCTCGCGCGCTCTTTGGCGCGGGC 604
DB 176 AlaLeuThrArgValProSerAlaLeuIleAlaLeuAlaProLeuLeuPheGlyArgGly 195
QY 605 GAGTGTGCGAGCTCGGCTCCAGCGTCCAGTGGAGCGGGAAACCTCTCTATGCGGCC 664
DB 196 GluValCysAspAlaArgLeuGlnArgCysGlnValSerArgGluProSerTyrAlaAla 215
QY 665 TTCTCCACCGCGCGCTCCACCTCGCGTGGCGTGGCGTGGCTGTGCTACCGGAAG 724
DB 216 PheSerThrArgGlyAlaPheHisLeuProLeuGlyValValProPheValTyrArgLys 235
QY 725 ATCTACGAGCGCGCAAGTTCGTTTCGCGCGCGCGGAGCTGTGCTGCGCTCCG 784
DB 236 IleTyrGluAlaAlaLysPheArgPheGlyArgArgArgAlaValLeuProLeuPro 255
QY 785 GCACCATGAGTGTGAGGTCAAGTAAAGAAAGCACTGATGAGGTGAGTGAAGTGTTC 844
DB 256 AlaThrMetGlnVal-----LysValLysGluAlaProAspGluAlaGluValValPhe 273
QY 845 ACGGCATTTGCAAGCAACGCTGCTCTCCAGTGGAGCGGGACTCTGCGCGGAGCAG 904
DB 274 ThrAlaHisCysLysAlaThrValSerPheGlnValSerGlyAspSerTrpArgGluGln 293
QY 905 AAGGAGAGCGAGCAGCATGATGGTGGGAATTCGATTGCGGTGTGTGTGCTGTGCG 964
DB 294 LysGluArgArgAlaAlaMetMetValGlyIleLeuIleGlyValPheValLeuCysTrp 313
QY 965 ATCCCTCTCTCTGAGGAACTCATCAGCCCACTGTGCTGCTGCTGCTGCTGCTGCTG 1024
DB 314 IleProPhePheLeuThrGluLeuLeuSerProLeuCysAlaCysSerLeuProPhe 333
QY 1025 TGAAGAAGCATATTCGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1084
DB 334 TrpLysSerIlePheLeuTrpLeuGlyTyrSerAsnSerPhePheAsnProLeuIleTyr 353
QY 1085 ACAGCTTTTAAAGAACTCAACAATGCTTCAAGAGCCTCTTACTAAGCAGAGA 1141
DB 354 ThrAlaPheAsnLysAsnTyrAsnAsnAlaPheLysSerLeuPheThrLysGlnArg 372
RESULT 6
ABG60235
ID ABG60235 standard; protein; 380 AA.
XX AC ABG60235;
XX DT 30-JUL-2002 (first entry)
XX

Human hydroxytryptamine receptor-like protein NOV8.
Human; NOVX; developmental disorder; endocrine disorder;
vascular disorder; infectious disease; anorexia; cancer; stroke;
neurodegenerative disorder; Alzheimer's disease; acute brain injury;
central nervous system disorder; depression; lung disorder;
reproductive disorder; tissue disorder; thrombocytopaenia; migraine;
angiogenesis; asthma; X-linked severe combined immunodeficiency;
inflammation; autoimmune disorder; immune disorder; blood disorder;
haematopoietic disorder; gastrointestinal disease; respiratory disorder;
hepatitis; fertility; hypertension; arteriosclerosis; ischaemia;
rheumatoid arthritis; Grave's disease; wound healing.

Homo sapiens.

WO200224733-A2.

28-MAR-2002.

17-SEP-2001; 2001WO-US029115.

15-SEP-2000; 2000US-0232675P.

15-SEP-2000; 2000US-0232676P.

15-SEP-2000; 2000US-0232679P.

18-SEP-2000; 2000US-0233382P.

18-SEP-2000; 2000US-0233402P.

19-SEP-2000; 2000US-0233521P.

19-SEP-2000; 2000US-0233522P.

19-SEP-2000; 2000US-0233801P.

20-SEP-2000; 2000US-0233960P.

06-OCT-2000; 2000US-0238398P.

13-OCT-2000; 2000US-0240284P.

13-OCT-2000; 2000US-0240498P.

11-JAN-2001; 2001US-0260973P.

26-JAN-2001; 2001US-0264274P.

09-MAR-2001; 2001US-0274862P.

(CURA-) CURAGEN CORP.

Mishra VS, Syptek KA, Taupier RJ, Vernet CAM, Colman SD;
Gorman L, Tchernev VT, Malyankar UM, Shenoy S, Tchernev VT;
Padigar M, Patturajan M, Burgess CE, Smithson G, Millet I;
Peyman JA, Stone D, Gunther E, Ellerman K;

WPI: 2002-383182/41.

N-PSDB; ABK1923.

New cytoplasmic, nuclear, membrane bound and secreted NOVX polypeptides,
useful for treating cancers and tumors, lung disorders, hematopoietic
disorders, autoimmune diseases and immune disorders.

Claim 1 ; Page 60; 210pp; English.

The invention relates to an isolated NOVX polypeptide selected from
NOV1a, NOV1b, NOV2a, NOV2b, NOV2c, NOV3a, NOV3b, NOV4a, NOV4b,
NOV5a, NOV5b or NOV6-NOV9 polypeptides, their mature form or variant.
Also included are a nucleic acid encoding a NOVX protein or variant; a
vector comprising the nucleic acid; a cell comprising the vector; an anti
-NOVX antibody; and identifying agents that modulate the expression or
activity of NOVX. NOVX, the nucleic acid, antibody and modulators are
useful in the diagnosis, treatment or prevention of developmental
disorders, endocrine disorders, vascular disorders, infectious disease,
anorexia, cancer, neurodegenerative disorders (e.g. Alzheimer's disease,
Parkinson's disease, Huntington's disease, multiple sclerosis and
amyotrophic lateral sclerosis), acute brain injury (e.g. stroke, head
injury and cerebral palsy), central nervous system disorders (e.g.
depression, epilepsy and schizophrenia), lung disorders, reproductive
disorders, disorders affecting carbohydrate metabolism (e.g.
galactosaemia and hereditary fructose intolerance), tissue disorders
(e.g. Wiskott-Aldrich syndrome, thrombocytopaenia, night blindness and
Pick's disease), disorders linked to abnormal angiogenesis, asthma,
azoospermia, learning disabilities, facial dysmorphism, autoimmune
encephalomyelitis, X-linked severe combined immunodeficiency, seizures,

CC migraines, inflammation, autoimmune disorders, disorders affecting sleep,
 CC appetite, thermoregulation, pain, perception, hormone secretion and
 CC sexual behaviour, immune disorders, haematopoietic disorders or other
 CC disorders related to cell signal processing and metabolic pathway
 CC modulation, gastrointestinal diseases, respiratory disorders, blood
 CC disorders, hepatitis, trauma, regeneration, viral, bacterial or parasitic
 CC infections, hyper- or hypo-thyroidism, endometriosis, fertility,
 CC hypertension, arteriosclerosis, ischaemia, haemolytic anaemia, Werner
 CC syndrome, rheumatoid arthritis, Grave's disease, wound healing, X-linked
 CC mental retardation, psychotic and neurological disorders and neuronal
 CC degeneration. The present sequence represents a NOVX protein
 XX
 SQ Sequence 380 AA;

Alignment Scores:

Pred. No.: 3,51e-133 Length: 380
 Score: 1819.50 Matches: 367
 Percent Similarity: 93.62% Conservative: 0
 Best Local Similarity: 93.62% Mismatches: 25
 Query Match: 84.47% Indels: 5
 DB: 5 Gaps: 5

US-09-976-782-15 (1-1152) x ABG60235 (1-380)

QY 5 ATGGAGGCGCTAGCTTTCAGTGGCCACCGCGCGGTTCCTTCCTCGCCGACCCGAG 64
 DB 1 MetGluAlaAlaSerLeuSerValAlaThrAlaGlyValAlaLeuAla-----ProGlu 18
 QY 65 ACCAGCAGC-----GGACCGGGACC 85
 DB 19 ThrSerProAlaLeuProLeuProTrpAspProArgProAlaAlaGlyProGlyThr 38
 QY 86 CCAACCCGAGGAGTACTCGTTCGACCCGCGCGCGCGTCTCGCCGCGCGAGGG 145
 DB 39 ProSerProArgGlyLeuGlySerThrProSerGlyAlaValLeuProGlyArgGly 58
 QY 146 CGGCCTCTCTGCTTTCAGGTCCTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 205
 DB 59 ProProPheSerValPheThrValLeuValValThrLeuLeuValLeuLeuAla 78
 QY 206 ACTTTCCTGTGAACCTGCTGTTCCGTCACATCCCGCGGGTCCGTGCTTCCACGCG 265
 DB 79 ThrPheLeuTrpAsnLeuLeuValProValThrLeuProArgValArgAlaPheHisArg 98
 QY 266 GTGCGCGATAACTGTCGCTCGCGCGCGTCTCGAGCACTAGTGGCAGCGTGGGG 325
 DB 99 ValProHisAsnLeuValAlaSerThrAlaValSerAspGluLeuValAlaAlaLeu 118
 QY 326 ATGCCACCGAGCTGCGGAGTGGTGTGCGACCGCGGCGAGTGGTGGTGGCGGAGC 385
 DB 119 MetProSerLeuAlaSerGluLeuSerThrGlyArgArgGluLeuGlyArg--- 137
 QY 386 CTGTGCGCAGTGTGATCTCTTGGACCGCGGAGCTGTCTGTGTGCGCCCGCGGCTC 445
 DB 138 -----HisValTrpIleSerPheAspAla-----LeuCysCysProAlaGlyLeu 152
 QY 446 GGGACGTCGCGCGCATCCGCTGGCGCGCGCGCATCACAGCGCACCTGCGAGCAC 505
 DB 153 GlyAsnValAlaAlaAlaLeuGlyArgAspGlyAlaIleThrArgHisLeuGlnHis 172
 QY 506 AGCGTCGCGACCGCGCGCGCTCGTGTGCTCATGTCGCGCTCGCCGCGGTGGCGTCG 565
 DB 173 ThrLeuArgThrArgSerArgAlaSerLeuLeuMetIleAlaLeuAlaArgValProSer 192
 QY 566 GGGCTCATCGCTCGCGCGCTGCTCTTGGCGGGGCGAGTGGCGACGCTCGGCTC 625
 DB 193 AlaLeuIleAlaLeuAlaProLeuLeuPheGlyArgGlyGluValCysAspAlaArgLeu 212
 QY 626 CAGCGCTGCGAGTGGCGGAAACCTCTATGCGCGCTTCTCCACCGCGCGGCTTC 685
 DB 213 GlnArgCysGlnValSerArgGluProSerTyAlaAlaPheSerThrArgGlyAlaPhe 232
 QY 686 CACCTGCGGCTTGGCGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 745

DB 233 HisLeuProLeuGlyValValProPheValTyArgLysIleTyGluAlaAlaLysPhe 252
 QY 746 GCTTTCGCGCCCGCGCGAGAGCTGTGCTGCGGTGGCGCCACCATCATGAGTGGAGTCC 805
 DB 253 ArgPheGlyArgArgArgAlaValLeuProLeuProAlaThrMetGln----- 269
 QY 806 AAGGTAAAGGAAGACACCTGATGAGGCTCAAGTGGTGTTCACGGCACATTCGAAAGCAAG 865
 DB 270 ---ValLysGluAlaProAspGluAlaGluValValPheThrAlaHisCysLysAlaThr 288
 QY 866 GTGTCTCTTCAGGTGAGCGGGACTCTCTGCGGGAGCAGAGGAGGAGGAGGAGGAGGAGG 925
 DB 289 ValSerPheGlnValSerGlyAspSerTrpArgGluGlnLysGluArgAlaMet 308
 QY 926 ATGTGGGAATCTGATTGGCGGTGTTGTGTGCTGCTGATCCCTCTTCTCTGACGAA 985
 DB 309 MetValGlyIleLeuIleGlyValPheValLeuGlyTrpIleProPheLeuThrGlu 328
 QY 986 CTCATCAGCCCACTCTGTGCTGCGAGCTGCCCCCATCTCGAAAGCATATTTCTGTGG 1045
 DB 329 LeuIleSerProLeuCysAlaCysSerLeuProProIleTrpLysSerIlePheLeuTrp 348
 QY 1046 CTTGCTACTCAATCTTCTTCAACCCCTGANTTACAGCTTTTAAAGAACTAC 1105
 DB 349 LeuGlyTyTrSerAsnSerPhePheAsnProLeuIleTyThrAlaPheAsnLysAsnTy 368
 QY 1106 AACATGCTTCAAGAGCTCTTTTACTAAGCAGAGA 1141
 DB 369 AsnAsnAlaPheLysSerLeuPheThrLysGlnArg 380
 RESULT 7
 AAR58686
 ID AAR58686 standard; protein; 370 AA.
 XX AAR58686;
 XX 25-MAR-2003 (revised)
 DT 24-MAY-1995 (first entry)
 XX Rat MR22 serotonin receptor protein.
 DE Serotonin; receptor; transmembrane; domain; kinase; phosphorylation;
 KW sensory; motor; behaviour; central nervous system; CNS; superfamily;
 KW G-protein; ligand-gated; ion channel; subfamily; human; rat; amplify;
 KW primer; PCR; amplification; brain; hypothalamus; indolamine; drug;
 KW hypothalamus; therapeutic; neurological; pathology; dementia; insomnia;
 KW Parkinson's disease; eating disorder; anxiety; migraine; headache.
 XX Rattus rattus.
 OS
 XX
 XX Key Location/Qualifiers
 FT Modified-site 5
 FT Domain /label= N-linked glycosylation
 FT Domain 53..78
 FT Domain /label= transmembrane domain
 FT Domain 89..115
 FT Domain /label= transmembrane domain
 FT Domain 126..149
 FT Modified-site 163
 FT Modified-site /label= kinase C phosphorylation site
 FT Domain 166
 FT Domain /label= kinase C phosphorylation site
 FT Domain 170..196
 FT Domain /label= transmembrane domain
 FT Domain 211..237
 FT Modified-site /label= transmembrane domain
 FT Modified-site 272
 FT Modified-site /label= kinase C phosphorylation site
 FT Modified-site 287
 FT Domain /label= kinase C phosphorylation site
 FT Domain 298..324


```
QY 1085 ACAGCTTTTACAAAGAACTACAAATGCGCTTCAAGAGCCTCTTTACTAAGCAGAGA 1141
DB :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
352 ThrAlaPheAsnLysAsnTyrAsnAsnAlaPheLysSerLeuPheThrLysGlnArg 370

RESULT 9
ID AAR57066
AA AAR57066 standard; protein; 370 AA.
XX
AC AAR57066;
XX
XX 25-MAR-2003 (revised)
DT 27-MAR-1995 (first entry)
XX
XX Murine serotoninergic receptor 5HT5b.
XX
XX Serotoninergic receptor 5HT5b; 5-hydroxytryptamine; neuromodulator;
KW mouse serotonin receptor; neurological; cardiovascular; psychiatric;
KW disorder.
XX
XX Mus musculus.
OS
XX
XX W09418319-A1.
PN
XX
XX 18-AUG-1994.
PD
XX
XX 07-FEB-1994; 94WO-FR000136.
PF
XX
XX 09-FEB-1993; 93FR-00001392.
PR
XX
XX (INRM ) INST NAT SANTE & RECH MEDICALE.
PA
XX
XX Amlaiky N, Boschert U, Grailhe R, Hen R, Matthes H, Plassat J;
PI
XX
XX WPI; 1994-279740/34.
DR
XX
XX N-PSDB; AAQ70264.
DR
XX
XX New serotoninergic receptor 5HT5b and related nucleic acid and
PT recombinant cells - useful for treatment and diagnosis of e.g.
PT neurological or cardiovascular disease, also for identifying serotonin
PT agonists and antagonists.
XX
XX Claim 1; Page 17-18; 27pp; French.
XX
XX Oligonucleotides corresponding to conserved regions of known serotonin
CC receptors were used in PCR of mouse brain RNA in presence of reverse
CC transcriptase and the products were sequenced. One product with homology
CC to known receptors was labelled and used to probe a cDNA library in
CC Lambda UniZAP. The 2036 bp sequence AAQ70264 was identified and deduced
CC to encode a protein having 7 hydrophobic domains which represents a novel
CC serotonin receptor which was designated 5HT5b (AAR57066). (Updated on 25-
CC MAR-2003 to correct PN field.) (Updated on 25-MAR-2003 to correct PA
CC field.)
XX
XX SQ Sequence 370 AA;

Alignment Scores:
Pred. No.: 1.36e-104 Length: 370
Score: 1452.50 Matches: 298
Percent Similarity: 83.64% Conservative: 19
Best Local Similarity: 78.63% Mismatches: 53
Query Match: 67.43% Indels: 9
DB: 2 Gaps: 3

US-09-976-782-15 (1-1152) x AAR57066 (1-370)
QY 5 ATGGAGGCGGTAGCTTTCAGTGGCCACCGCGCGGTGGCTTGCCTCGCCCTGGAGCCCGAG 64
DB :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1 MetGluValSerAsnLeuSerGlyAlaThrProGlyLeuAlaPheProGlyProGlu 20

65 ACCAGCAGCGGACCGGAGCCCAAGCCGAGAGGATACCTCGGTTCGACCCCGAGCGGC 124
DB :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
21 SerCysSer-----AspSerProSerSerGlyArgSerMetGlySerThrProGlyGly 38
```

```
QY 125 GCCGTCTCGCGGCGCGAGGCGCGCCTTCTCTCTTACGGTCTCTGGTGTGACGCTG 184
DB :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
39 LeuIleLeuProGlyArgGluProProPheSerAlaPheThrValLeuValValThrLeu 58

185 CTAGTGTGTGTGATCGCTGCCACTTCTCTGTGGAACCTGTGTGGTTCGGTCAACATCCG 244
DB :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
59 LeuValLeuLeuIleAlaAlaThrPheLeuTrpAsnLeuLeuValLeuValThrIleLeu 78

245 CGGTCCTCGTCTCCACCGCGTCCGATAACTTGGTGGCCTCGACGGCGCTCCGGAC 304
DB :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
79 ArgValArgAlaPheHisArgValProHisAsnLeuValAlaSerThrAlaValSerAsp 98

305 GACTAGTGGCAGCGCTGGCGATGCCACCGAGCTGGCGAGTGGTGTCTGACCGCGGCGA 364
DB :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
99 ValLeuValAlaValLeuValMetProLeuSerLeuValSerGluLeuSerAlaGlyArg 118

365 CGTCGGCTGTGGCGCGAGCCTGTGCCAGTGTGATCTCTTCGACGGCGGAGCCTGT 424
DB :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
119 ArgTrpGlnLeuGlyArgSerLeuCysHisValTrpIleSerPheAspVal----- 135

425 CTGTGCTGCCCGCGCGCTCGGGAACGTGGCGCGCATCGCCCTGGCGCGACGGGGCC 484
DB :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
136 LeuCysCysThrAlaSerIleTrpAsnValAlaAlaIleAlaLeuAspArgTyrTrpThr 155

485 ATCACAGCGCACCTGCAGCACACCTGGCAACCGCGCGCGCTCGTCTCTCATGATC 544
DB :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
156 IleThrArgHisLeuGlnTyrThrLeuArgThrArgSerArgAlaSerAlaLeuMetIle 175

545 GCGCTCGCCCGGGTCCGCTCATCGCCCTCGCGCGCTGCTCTTTGGCGGGGC 604
DB :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
176 AlaIleThrTrpAlaLeuSerAlaLeuIleAlaLeuAlaProLeuLeuPheGlyTrpGly 195

605 GAGTGTGCGACGCTCGGCTCCAGCGTCCAGGTGAGCGCGGGAACCTCTATGCGGCC 664
DB :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
196 GluAlaTyrAspAlaArgLeuGlnArgCysGlnValSerGlnGluProSerTyrAlaVal 215

665 TTCTCCACCGCGCGCTTCCACCTCGCGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 724
DB :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
216 PheSerThrCysGlyAlaPheTyrIleProLeuAlaValValLeuPheValTyrTrpLys 235

725 ATCTACGAGCGGCCAAAGTTTCGTTTCGCGCGCGCGCGAGAGTGTGCTCGGTGCGG 784
DB :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
236 IleTyrLysAlaAlaLysPheArgPheGlyArgArgArgAlaValValProLeuPro 255

785 GCCACCATGAGGTGAGTCCAGTAAAGAACACCTGATGAGGTGAAGTGGTGTTC 844
DB :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
256 AlaThrThrGln-----AlaLysGluAlaProProGluSerGluMetValPhe 271

845 ACGGCACATTGCAAGCAACCGTCTCTCCAGGTGAGCGGGGACTCTCTGGCGGAGCAG 904
DB :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
272 ThrAlaArgArgAlaThrValThrPheGlnThrSerGlyAspSerTrpArgGluGln 291

905 AAGGAGAGCGGAGCAGCATGATGGTGGGAATCTGATTGGCGTGGTGGTGGTGGTGG 964
DB :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
292 LysGluLysArgAlaAlaMetMetValGlyIleLeuIleGlyValPheValLeuCysTrp 311

965 ATCCCTCTCTTCGAGGAACTCATCAGCCCATCTGTGCTGCTGAGCCTGCCCCCATC 1024
DB :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
312 IleProPhePheLeuThrGluLeuIleSerProLeuCysAlaCysSerLeuProProlle 331

1025 TGGAAAAGCATATTCTGTGGCTGGCTACTCCAAATCTTTCTTCAACCCCTCATTTAC 1084
DB :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
332 TrpLysSerIlePheLeuTrpLeuGlyTyrSerAsnSerPhePheAsnProLeuIleTyr 351

1085 ACAGCTTTTAAAGAACTACAAATGCGCTTCAAGAGCCTCTTTTACTAAGCAGAGA 1141
DB :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
352 ThrAlaPheAsnLysAsnTyrAsnAsnAlaPheLysSerLeuPheThrLysGlnArg 370

RESULT 10
AAR57066
ID AAR57066 standard; protein; 370 AA.
XX
```

AC AAE31673;
 XX 07-MAR-2003 (first entry)
 XX Mouse 5-HT5B receptor protein.
 DE Mouse; 5-hydroxytryptophan; 5-HT5B receptor; depression; transgenic;
 KW transgenic animal; phenotype; pain sensitivity; neurological disease;
 KW gene therapy; pharmacological; neuropsychological disease; analgesic;
 KW psychotic illness; receptor.
 XX Mus sp.
 XX WO200279443-A2.
 XX 10-OCT-2002.
 XX 29-MAR-2002; 2002WO-US009853.
 XX 29-MAR-2001; 2001US-0280553P.
 XX 21-DEC-2001; 2001US-0342472P.
 XX 28-MAR-2002; 2002US-00109532.
 XX (DELT-) DELTAGEN INC.
 XX Allen KD;
 XX WPI; 2003-067450/06.
 XX N-PSDB; AAD48747.
 XX Novel transgenic mouse useful for identifying a potential therapeutic
 PT agent for the treatment of depression or pain, comprises disruption in a
 PT 5-hydroxytryptophan 5B receptor gene.
 XX Disclosure; Fig 2; 60pp; English.
 XX The present invention relates to transgenic animals, compositions and
 CC methods relating to the characterisation of gene function. The invention
 CC also relates to transgenic mice comprising mutations in 5-hydroxy-
 CC tryptophan (5-HT)5B receptor gene. The transgenic mice are useful for
 CC identifying an agent that modulates the phenotype such as increased
 CC depression or increased pain sensitivity. They are also useful for
 CC identifying potential therapeutic agents for the treatment of pain or
 CC depression and for evaluating a potential therapeutic agent capable of
 CC affecting a condition associated with a mutation in a 5-HT5B receptor
 CC gene. Transgenic animals of the invention are also useful for testing the
 CC efficacy of proposed genetic and pharmacological therapies for human
 CC diseases such as neurological, neuropsychological or psychotic illnesses.
 CC The present sequence is mouse 5-HT5B receptor protein
 XX Sequence 370 AA;
 SQ
 Alignment Scores:
 Pred. No.: 1.36e-104 Length: 370
 Score: 1452.50 Matches: 298
 Percent Similarity: 83.64% Conservative: 19
 Best Local Similarity: 78.63% Mismatches: 53
 Query Match: 67.43% Indels: 9
 DB: 6 Gaps: 3
 US-09-976-782-15 (1-1152) x AAE31673 (1-370)
 QY 5 ATGAGGCGGTAGCTTTTCAGTGGCCACCGCGGCTTGCCTTCCCTGGGACCGGAG 64
 DB 1 MetGluValSerAsnLeuSerGlyAlaThrProGlyLeuAlaPheProGlyProGlu 20
 QY 65 ACCAGCAGCGGACCGGGACCCAAAGCCCGAGAGGGATACCTCGTTTCGACCCCGCGGC 124
 DB 21 SerCysSer-----AspSerProSerGlyArgSerMetGlySerThrProGlyGly 38
 QY 125 GCGTCTCTGCGGCGGCGGCGCGCTTCTGCTTCAACCGCTCTGCTGCTGCTGCTGCTG 184
 DB 39 LeuIleLeuProGlyArgGluProProPheSerAlaPheThrValLeuValThrLeu 58

QY 185 CTAGTGTGCTGATCGTCCCACTTTCTGTGTGAACCTGCTGTGTTCCGGTCCACCTCCCG 244
 DB 59 LeuValLeuLeuIleAlaAlaThrPheLeuTrpAsnLeuValLeuValThrIleLeu 78
 QY 245 CCGGTCCGTGCTTCACCGCGTCCGATACCTTGGTGGCTCCGACGCGCTCTCGGAC 304
 DB 79 ArgValArgAlaPheHisArgValProHisAsnLeuValAlaSerThrAlaValSerAsp 98
 QY 305 GAACCTAGTGGCAGCGCTGCGATGCCACCGACCTGGCGAGTGTGCTGTCGACCGGGCGA 364
 DB 99 ValLeuValAlaValLeuValMetProLeuSerLeuValSerGluLeuSerAlaGlyArg 118
 QY 365 CGTGTGCTGCTGGCGCGGAGCTGTGCCACGTGTGATCTCTCTGTGACCGCGGAGCTGT 424
 DB 119 ArgTrpGlnLeuGlyArgSerLeuCysHisValTrpIleSerPheAspVal----- 135
 QY 425 CTGTGTGCTCCCGCGGCTCTCGGGAACGTGGCGGCATCGCCCTGGCGCGACGGGGCC 484
 DB 136 LeuCysCysThrAlaSerIleTrpAsnValAlaAlaIleAlaLeuAspArgTyrTrpThr 155
 QY 485 ATCACACGGCAGCTGACGACAGCTGGCGCACCGCGAGCGCGCTCTGCTCATGATC 544
 DB 156 IleThrArgHisLeuGlnTyrThrLeuArgThrArgSerArgAlaSerAlaLeuMetIle 175
 QY 545 GCGTCTGCGCGCGGTGCGCTCATCGCCCTCGCGCCCTGCTCTTTGGCGGGGC 604
 DB 176 AlaIleThrTrpAlaLeuSerAlaLeuIleAlaLeuAlaProleuLeuPheGlyTrpGly 195
 QY 605 GAGGTGTGCGACGCTCGCTCCAGCGCTGCCAGGTGAGCGCGGAACCTCTATGCGGCC 664
 DB 196 GluAlaTyrAspAlaArgLeuGlnArgCysGlnValSerGlnGluProSerTyrAlaVal 215
 QY 665 TTCCTCACCGCGCGGCTTCCACTGCGCTTGGCGTGGTGGTGGTGGTGGTGGTGGTGG 724
 DB 216 PheSerThrCysGlyAlaPheTyrLeuProLeuAlaValLeuPheValTyrTrpLys 235
 QY 725 ATCTACGAGCGCGCAAGTTTCGTTTGGCGCGCGCGGAGAGCTGTGCTGCGTGGCG 784
 DB 236 IleTyrLysAlaAlaLysPheArgPheGlyArgArgArgAlaValValProleuPro 255
 QY 785 GCCACCATGCGAGTGTGAGTTCGAAGTAAAGGAAGCACTGTAGTGTGAGTGTGTTTC 844
 DB 256 AlaThrThrGln-----AlaLysGluAlaProProGluSerGluMetValPhe 271
 QY 845 ACGGCACATTCGAACCAACGCTGCTTCCAGGTGAGCGGGGACTCCTGGCGGGAGAG 904
 DB 272 ThrAlaArgArgAlaThrValThrPheGlnThrSerGlyAspSerTrpArgGluGln 291
 QY 905 AAGGAGAGCGGAGCAGCATGATGCTGGGAATTCGATTGGCGTGTGCTGTGCTGTG 964
 DB 292 LysGluLysArgAlaAlaMetMetValGlyIleLeuIleGlyValPheValLeuCysTrp 311
 QY 965 ATCCCTCTTCTCCACGGAACCTATACGCCCACTGTGTGCTGAGCGCTGCCCCCATC 1024
 DB 312 IleProPhePheLeuThrGluLeuIleSerProLeuCysAlaCysSerLeuProIle 331
 QY 1025 TGGAAAGCATATTTCTGTGGCTTGGCTACTCCAATTCTTTCTTCAACCCCTGATTAC 1084
 DB 332 TrpLysSerIlePheLeuTrpLeuGlyTyrSerAsnSerPhePheAsnProleuIleTyr 351
 QY 1085 ACAGCTTTTAAACAAGAACTACAACAATGCTTCAAGAGCTCTTACTAAGCAGAGA 1141
 DB 352 ThrAlaPheAsnLysAsnTyrAsnAsnAlaPheLysSerLeuPheThrLysGlnArg 370
 RESULT 11
 ADC86157
 ID ADC86157 standard; protein; 606 AA.
 XX
 AC ADC86157;
 XX
 DT 01-JAN-2004 (first entry)
 XX

DE Human GPCR protein SEQ ID NO:610.
 KW human; GPCR; guanosine triphosphate-binding protein coupled receptor;
 KW gene therapy.
 OS Homo sapiens.
 XX EPI270724-A2.
 PN 02-JAN-2003.
 XX 18-JUN-2002; 2002EP-00013517.
 PF 18-JUN-2001; 2001JP-00246789.
 PR (NAAD-) NAT INST ADVANCED IND SCI & TECHNOLOGY.
 XX (ADSC-) CENT ADVANCED SCI & TECHNOLOGY INCUBATIO.
 PA Suwa M, Asai K, Akiyama Y, Aburatani H;
 FI WPI; 2003-315783/31.
 DR N-PSDB; ADC86156.
 XX New polynucleotide, useful for preparing a composition for treating a
 PT patient in need of increased or suppressed activity or expression of the
 PT guanosine triphosphate-binding protein coupled receptor.
 XX Claim 2; SEQ ID NO 610; 28pp; English.
 CC The invention relates to a novel polynucleotide encoding a guanosine
 CC triphosphate-binding protein coupled receptor (GPCR). A polynucleotide of
 CC the invention may have a use in gene therapy. The polynucleotide and
 CC polypeptide are useful for preparing a composition for treating a patient
 CC in need of increased or suppressed activity or expression of the
 CC guanosine triphosphate-binding protein coupled receptor. The protein
 CC sequences shown in ADC8549-ADC87617 represent GPCR's of the invention.
 XX
 SQ Sequence 606 AA;
 Alignment Scores:
 Pred. No.: 2,16e-92 Length: 606
 Score: 1296.50 Matches: 274
 Percent Similarity: 83.33% Conservative: 6
 Best Local Similarity: 81.55% Mismatches: 24
 Query Match: 60.19% Indels: 33
 DB: 7 Gaps: 5
 US-09-976-782-15 (1-1152) x ADC86157 (1-606)
 QY 3 CCATGAGGCGCTAGCTTTTCAGTGGCCACGCGCGGTTGCCCTTGGCCCTGGGACCCG 62
 Db 254 ProTprargProLeuAlaPheGlnTrpProProAlaLeuProLeuProTpraspPro 273
 QY 63 AGACCAGGAGC-GGACCGGACCCCAAGCCGAGAGGATACCTCGGTTCCACCCGAGC 121
 Db 274 ArgProAlaAlaGlyProGlyThrProSerProargGlyIleLeuGlySerThrProSer 293
 QY 122 GCGCGCTCTCTGCGCGCGGAGGCGCCCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 181
 Db 294 GlyAlaValLeuProGlyArgGlyProProPheSerValPheThrValLeuValThr 313
 QY 182 CTGCTAGTGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 241
 Db 314 LeuLeuValLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeu 333
 QY 242 CCGCGGTCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 301
 Db 334 ProArgValArgAlaPheHisArgValProHisAsnLeuValAlaSerThrAlaValSer 353
 QY 302 GACGAACCTAGTGGAGCGCTGGCGATGCCACCGAGCTGGCGAGTGGAGTGGAGTGG 361
 Db 354 AspGluLeuValAlaAlaLeuAlaMetProProSerLeuAlaSerGluLeuSerThrGly 373

QY 362 CGACGTGGCTGCTGGCGCGAGCTGTGCCACGTGTGGATCTCTCTCGACCGGAGCC 421
 Db 374 ArgArgArgLeuLeuGlyArgSerLeuCysHisProValProArgValAspLeuLeuPro 393
 QY 422 TGTCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 481
 Db 394 ArgLeuCysCysProAlaGlyLeuGlyAsnValAlaAlaLeuGlyArgAspGly 413
 QY 482 GCCATCACAGCGCAGCTGACACACACGCTGGCCACCGCGCGCTGCTGCTGCTGCTGCTG 541
 Db 414 AlaIleThrArgHisLeuGlnHisThrLeuArgThrArgSerArgAlaSerLeuLeuMet 433
 QY 542 ATCGCGCTGCGCGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 601
 Db 434 IleAlaLeuAlaArgValProSerAlaLeuLeuAlaLeuAlaProLeuLeuPheGlyArg 453
 QY 602 GCGAGGTGTGCGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 661
 Db 454 GlyGluValCysAspAlaArgLeuGlnArgCysGlnValSerArgGluProSerTyrAla 473
 QY 662 GCCTTCTCCACCGCGCGGCTTCCACCTGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 721
 Db 474 AlaPheSerThrArgGlyAlaPheHisLeuProLeuGlyValValProPheValTyrArg 493
 QY 722 AAGATCTACAGAGCGCGCAAGTTTCGTTTCGCGCGCGCGCGAGAGCTGTGCTGCGCTG 781
 Db 494 LysIleTyrGluAlaAlaLysPheArgPheGlyArgArgAlaValLeuProLeu 513
 QY 782 CCGGCCACCATGCGAGGTGAGGTCCAAAGGTA----- 811
 Db 514 ProAlaThrMetGlnValArgGlyGlyLeuArgAsnValAlaMetGlyLysArgLeuLeu 533
 QY 812 ---RAGGAAGCACCTGATGAGGCTG---AAGTGG-----TGTTCACGGCAGCATTCG 856
 Db 534 GlulysGluAlaAlaSer-ArgMetGlyGluTrpAlaGluAlaCysThrAsnGlyAlaAr 553
 QY 857 AAAGCAACGCGTGTCT-----TCCAGGTGAGCGGG 886
 Db 553 gAlaGlnArgSerProGlyAlaHisGluAspLysPheAlaIleSerSerGluAlaGln 573
 QY 887 GACTCTCTGCG-----GGGAGCAGAGGAGAGGC 914
 Db 573 yThrGluGlyLeuValThrGlySerProGlyThrGlnValArgGly 588
 RESULT 12
 AAU79252
 ID AAU79252 standard; protein; 357 AA.
 AC AAU79252;
 DT 30-JUL-2002 (first entry)
 XX Human 5-hydroxytryptamine receptor 5A (HTR5A).
 DE Human; 5-hydroxytryptamine receptor 5A; HTR5A; serotonin;
 KW neuroprotective; neurological disease; depression; epilepsy;
 KW gene therapy; single nucleotide polymorphism; haplotype pair.
 OS Homo sapiens.
 XX WO200222887-A1.
 XX 21-MAR-2002.
 XX 17-SEP-2001; 2001WO-US029210.
 XX 15-SEP-2000; 2000US-0233051P.
 XX (GENA-) GENAISSANCE PHARM INC.
 XX Kazemi A, Koshy B, Sanchis A, Tirrell C;
 XX WPI; 2002-393978/42.
 DR

DR N-PSDB; ABK50434.

XX Novel genetic variants of 5-Hydroxytryptamine (Serotonin) Receptor 5A

PT isogenes, useful for improving efficiency and reliability in drug

PT development for treating neurological diseases.

XX

PS Claim 31; Fig 4; 134pp; English.

XX

CC The invention relates to single nucleotide polymorphisms in the gene

CC encoding human 5-hydroxytryptamine (serotonin) receptor 5A (HTR5A). A

CC method for haplotyping the HTR5A gene in an individual comprises

CC identifying the nucleotide at one or more polymorphic sites and

CC determining whether one of the copies of the gene is defined by one of

CC the HTR5A haplotypes given in the specification or whether both copies

CC are defined by a haplotype pair. This method is useful in genotyping,

CC whereby all possible haplotype pairs can be assigned to specific

CC genotypes. An association between a trait and a haplotype or haplotype

CC pair of the HTR5A gene can be identified by comparing the frequency of

CC the haplotype or haplotype pair in a population exhibiting the trait with

CC the frequency of the haplotype or haplotype pair in a reference

CC population, where a higher haplotype frequency in the trait population

CC indicates the trait is associated with the haplotype or haplotype pair.

CC HTR5A and its corresponding DNA are used for studying the expression and

CC function of HTR5A, and in screening for candidate drugs to treat diseases

CC related to HTR5A activity, such as neurological disorders, including

CC depression and epilepsy. This sequence represents the human HTR5A

CC polypeptide

XX

SQ Sequence 357 AA;

Alignment Scores:

Pred. No.:	2,73e-74	Length:	357
Score:	1063.50	Matches:	226
Percent Similarity:	72.98%	Conservative:	36
Best Local Similarity:	62.95%	Mismatches:	86
Query Match:	49.37%	Indels:	11
DB:	5	Gaps:	6

US-09-976-782-15 (1-1152) x RAU79252 (1-357)

QY 65 ACCAGACGGACCGCGGACCCCAAGCCGAGAGGATACCTCGTTGACCCCGACGGCG 124

DB 8 ThrSerPheSerLeuSerThrProSerProLeuGluThrAsnHisSerLeuGlyLysAsp 27

QY 125 GCGGCTCTCGGCGCGGCGCGCCCTCTCTGCTCTTACCGTCTGCTGCTGCTGCTG 184

DB 28 AspLeuArgProSer---SerProLeuLeuSerValPheGlyValLeuLeuLeuThrLeu 46

QY 185 CTAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 244

DB 47 LeuGlyPheLeuValAlaAlaThrPheAlaTrpAsnLeuLeuValLeuAlaThrLeuLeu 66

QY 245 CGGGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 304

DB 67 ArgValArgThrPheHisArgValProHisAsnLeuValAlaSerMetAlaValSerAsp 86

QY 305 GAACTAGTGCAGCGCTGCGGATGCCACCGAGCTGCGAGTGTGCTGCTGCTGCTGCTG 364

DB 87 ValLeuValAlaAlaLeuValMetProLeuSerLeuValHisGluLeuSer---GlyArg 105

QY 365 CGTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 424

DB 106 ArgTrpGlnLeuGlyArgLeuCysGlnLeuThrTrpLeuAlaCysAspVal----- 122

QY 425 CTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 484

DB 123 LeuCysCysThrAlaSerIleTrpAsnValThrAlaIleAlaLeuAspArgTyrTrpSer 142

QY 485 ATCACAGCGCACCTGCGAGCAGCTGCGCACCGCGAGCGCGCTGCTGCTGCTGCTGCTG 544

DB 143 IleThrArgHisMetGluTrpThrLeuArgThrArgLysCysValSerAsnValMetIle 162

QY 545 GCGCTGCGCGCGGCTGCGCTGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 604

DB 163 AlaLeuThrTrpAlaLeuSerAlaValIleSerLeuAlaProLeuLeuPheGlyTrpGly 182

QY 605 GAGGTGTGCGAGCTCGGCTCCAGCGCTGCGAGGTGAGCGCGGAGACCTCTCTATGCGGCC 664

DB 183 GluThrTyrSerGluGlySerGluCysGlnValSerArgGluProSerTyrAlaVal 202

QY 665 TTCTCCACCGCGCGCGCTTCCACCTGCGCTGCGTGGCTGGTGGCTGGTGGTGGTGGTGG 724

DB 203 PheSerThrValGlyAlaPheTyrLeuProLeuCysValValLeuPheValTyrTrpLys 222

QY 725 ATCTACGAGCGCGCAAGTTTCGTTTCGGC---CGCGCGCGGAGAGCTGCTGCTGCTGCTG 781

DB 223 IleTyrLysAlaAlaLysPheArgValGlySerArgLysThrAsnSerValSerProIle 242

QY 782 CCGGCGCACCATCGAGTGGTCCCAAGGTAAAGGAAGCACCTGATGAGGCTGAAGTGGTG 841

DB 243 SerGluAlaValGluValLysAspSerAlaLys-----GlnProGlnMetVal 258

QY 842 TTCACGCGCACATTGCAAGCAACGGTGTCTTCCAGGTGAGCGGGGACTCTCTGGCGGGAG 901

DB 259 PheThrValArg---HisAlaThrValThrPheGlnProGluGlyAspThrTrpArgGlu 277

QY 902 CAGAGGAGAGCGAGCGAGCCATGATGTGGGAATCTCTGATGGCGGTGTTGCTGCTGCTG 961

DB 278 GlnLysGluGlnArgAlaAlaLeuMetValGlyIleLeuIleGlyValPheValLeuCys 297

QY 962 TGGATCCCTCTCTCTGAGCGGAACCTCATCAGCCCACTCTGCTGCTGAGCGCTGCCCCC 1021

DB 298 TrpIleProPhePheLeuThrGluLeuIleSerProLeuCysSerCysAspIleProAla 317

QY 1022 ATCTGGAAGACATATTTCTGTGGTGTGGTGTGGTGTGGTGTGGTGTGGTGTGGTGTGG 1081

DB 318 IleTrpLysSerIlePheLeuTrpLeuGlyTyrSerAsnSerPhePheAsnProLeuIle 337

QY 1082 TACACAGCTTTTAAAGAACTACAACTGCTTCAAGAGCTCTTTTACTAGCAG 1138

DB 338 TyrThrAlaPheAsnLysAsnTyrAsnSerAlaPheLysAsnPhePheSerArgGln 356

RESULT 13

ABP81980

ID ABP81980 standard; protein; 357 AA.

XX

AC ABP81980;

XX

DT 04-MAR-2003 (first entry)

XX

DE Human 5-HT5A receptor protein SEQ ID NO:446.

XX

KW G protein-coupled receptor; GPCR; antigenic peptide; gene therapy;

KW G protein-coupled receptor; modulator; antibody; immune-related disease;

KW growth-related disease; cell regeneration-related disease; AIDS; cancer;

KW immunological-related cell proliferative disease; autoimmune disease;

KW Alzheimer's disease; atherosclerosis; infection; osteoarthritis; allergy;

KW osteoporosis; cardiomyopathy; inflammation; Crohn's disease; diabetes;

KW graft versus host disease; Parkinson's disease; multiple sclerosis; pain;

KW psoriasis; anxiety; depression; schizophrenia; dementia; memory loss;

KW mental retardation; epilepsy; asthma; tuberculosis; obesity; nausea;

KW hypertension; hypotension; renal disorder; rheumatoid arthritis; trauma;

KW ulcer.

XX

OS Homo sapiens.

XX

XX WO200261087-A2.

PN

XX

PD 08-AUG-2002.

XX

PF 19-DEC-2001; 2001WO-US050107.

XX

PR 19-DEC-2000; 2000US-0257144P.

XX

PA (LIFE-) LIFESPAN BIOSCIENCES INC.

XX

PR 18-JAN-2002; 2002US-0349511P.
 PR 28-FEB-2002; 2002US-0360500P.
 PR 15-MAR-2002; 2002US-0365041P.
 PR 19-APR-2002; 2002US-0374063P.
 PR 14-AUG-2002; 2002US-0403468P.
 PR 27-SEP-2002; 2002US-0414262P.
 PR 21-OCT-2002; 2002US-0419986P.
 PR 05-NOV-2002; 2002US-0423809P.
 PR 26-NOV-2002; 2002US-0429797P.
 XX
 PA (MILL-) MILLENNIUM PHARM INC.
 XX
 PI Silos-Santiago I, Karicheti V;
 XX
 DR WPI: 2003-598705/56.
 DR N-PSDB; AAD58496.
 XX
 PT Identifying a compound for treating urological disorders, for example
 PT urinary incontinence by assaying the ability of the compound to modulate
 PT the nucleic acid expression or polypeptide activity.
 XX
 PS Disclosure; Page 187; Opp; English.
 XX
 CC The present relates to a method for identifying a compound for treating
 CC urological disorders e.g., urinary incontinence including overactive/
 CC oversensitive bladder, overflow urinary incontinence, stress urinary
 CC incontinence caused by dysfunction of the bladder, urethra or central or
 CC peripheral nervous system, prostatitis, benign prostatic hyperplasia,
 CC cancer of the prostate or kidney disorders. The method is also useful for
 CC modulating hyperplasia in a cell and treating a subject having a
 CC urological disorder. The invention is also used in gene therapy. The
 CC present sequence is human GPCR known as 5-hydroxytryptamine 5A (5-HT-5A)
 CC (serotonin receptor)
 XX
 SQ Sequence 357 AA;

Alignment Scores:
 Pred. No.: 2,73e-74 Length: 357
 Score: 1063.50 Matches: 226
 Percent Similarity: 72.98% Conservative: 36
 Best Local Similarity: 62.95% Mismatches: 86
 Query Match: 49.37% Indels: 11
 DB: 7 Gaps: 6

US-09-976-782-15 (1-1152) x AA#38595 (1-357)

QY 65 ACCAGCAGCGACCGCGGACCCCAAGCCGAGGAGGATCTCGTTGACCCCGAGCGGC 124
 DB 8 ThrSerPheSerLeuSerThrProSerProLeuGluThrAsnHisSerLeuGlytysAsp 27
 QY 125 GCGGTCTCTCGCGGCGAGGCGCGCTTCTCTGCTTCCACGCTCTGTGGTGAGCGTG 184
 DB 28 AspLeuArgProSer---SerProLeuSerValPheGlyValleuLeuThrLeu 46
 QY 185 CTAGTGCTCTGCTGCTCCACTTTCCTGTGGAACCTGCTGCTCCGCTCCACATCCCG 244
 DB 47 LeuGlyPheLeuValAlaAlaThrPheAlaTrpAsnLeuValLeuAlaThrileLeu 66
 QY 245 CGGCTCGGTCCTTCCACCGCTGCGGATACCTTGGTGGCTCGAGCGGCTCTCGAC 304
 DB 67 ArgValArgThrPheHisArgValProHisAsnLeuValAlaSerMetAlaValSerAsp 86
 QY 305 GAACTAGTGGACGCTGGCGATGCCACCGAGCTGGCGAGTGAGCTGTCGACCGGCGCA 364
 DB 87 ValLeuValAlaAlaLeuValMetProLeuSerLeuValHisGluLeuSer---GlyArg 105
 QY 365 CGTGGCTCTGGCGGAGCCTGTGCCAGCTGTGGATCTCTTTCGACCGGAGCCTGT 424
 DB 106 ArgTrpGlnLeuGlyArgArgLeuGlyGlnLeuTrpIleAlaCysaspVal----- 122
 QY 425 CTGTGCTGCGCGCGCTCGGAGACGTGGCGCATCGCCCTGGCGCGGCGGCGGC 484
 DB 123 LeuCysCysThrAlaSerIleTrpAsnValThrAlaIleAlaLeuAspArgTyrTrpSer 142

QY 485 ATCACACGGCAGCTGCAGCACACACGCTGCACCGCAGCGCGCTCGTGTCTCATGATC 544
 DB 143 IleThrArgHisMetGluTyrThrLeuArgThrArgLysCysValSerAsnValMetIle 162
 QY 545 GCGCTGCGCCGGTGCCTGTCGCGCTCATGCGCCCTGCGCCGCTGTCTTTTGGCCGGGC 604
 DB 163 AlaLeuThrTrpAlaLeuSerAlaValIleSerLeuAlaProLeuLeuPheGlyTrpGly 182
 QY 605 GAGGTGTGCAGCGCTCGCGCTCCAGCGCTCCAGGTGAGCGGAGCGGACCTCTATGCGGC 664
 DB 183 GluThrTyrSerGluGlySerGluGluCysGlnValSerArgGluProSerTyrAlaVal 202
 QY 665 TTCTCCACCGCGCGCTTCCACCTGCGCTTGGCGTGTGGCGTGTGGCGTGTCTTACCGGAAG 724
 DB 203 PheSerThrValGlyAlaPheTyrLeuProLeuCysValValLeuPheValTyrTrpLys 222
 QY 725 ATCTACGAGCGCGCAAGTTTCGTTTCGCGC---CGCCGCGGAGAGCTGTGTCGCGTTG 781
 DB 223 IleTyrLysAlaAlaLysPheArgValGlySerArgLysThrAsnSerValSerProIle 242
 QY 782 CCGGCCACCATGCAGGTGAGTCCCAAGGTAAGGAGACCTGATGAGGCTGAAGTGTG 841
 DB 243 SerGluAlaValGluValLysAspSerAlaLys-----GlnProGlnMetVal 258
 QY 842 TTTCAGCGCACATTGCAAGCAACGGTGTCTTCCAGGTGAGCGGCGACTCTCGCGGAG 901
 DB 259 PheThrValArg--HisAlaThrValThrPheGlnProGluGlyAspThrTrpArgGlu 277
 QY 902 CAGAGGAGAGCGGAGCAGCAGCCATGATGTGGGAATTCCTGATGCGCTGTGTGCTGTGC 961
 DB 278 GlnLysGluGlnArgAlaAlaLeuMetValGlyIleLeuIleGlyValPheValLeuCys 297
 QY 962 TGGATCCCTCTCTCTGAGCGGAACCTATCAGCCCACTCTGTGCTGCGAGCTGCCCCC 1021
 DB 298 TrpIleProPhePheLeuThrGluLeuIleSerProLeuCysSerCysAspIleProAla 317
 QY 1022 ATCTGGAAGCATATTTCTGTGGCTTGGCTACTTCCAAATCTTCTTCAACCCCTGATT 1081
 DB 318 IleTrpLysSerIlePheLeuTrpLeuGlyTyrSerAsnSerPhePheAsnProLeuIle 337
 QY 1082 TACACAGCTTTTAAAGAACTACAACTGCTTCAAGAGCTCTTTACTAAGCAG 1138
 DB 338 TyrThrAlaPheAsnLysAsnTyrAsnSerAlaPheLysAsnPhePheSerArgGln 356
 RESULT 15
 ADE56996
 ID ADE56996 standard; protein; 357 AA.
 XX
 AC ADE56996;
 XX
 DT 29-JAN-2004 (first entry)
 XX
 DE Human Protein NP_076917, SEQ ID NO 2851.
 XX
 KW Human; pain; neuronal tissue; gene therapy;
 KW spinal segmental nerve injury; chronic constriction injury; CCI;
 KW spared nerve injury; SNI; Chung.
 XX
 OS Homo sapiens.
 XX
 PN WO2003016475-A2.
 XX
 PD 27-FEB-2003.
 XX
 PP 14-AUG-2002; 2002WO-US025765.
 XX
 PR 14-AUG-2001; 2001US-0312147P.
 PR 01-NOV-2001; 2001US-0346382P.
 PR 26-NOV-2001; 2001US-033347P.
 XX
 XX (GEHO) GEN HOSPITAL CORP.
 PA (FARB) BAYER AG.

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - protein search, using frame_plus_n2p model

Run on: August 31, 2004, 20:06:29 ; Search time 158 Seconds
(without alignments)
4600.974 Million cell updates/sec

Title: US-09-976-782-15

Perfect score: 2154

Sequence: 1 cgccatggagccgctagcc.....aagcagagatgaacacaggg 1152

Scoring table:

BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 1017041 seqs, 315518202 residues

Total number of hits satisfying chosen parameters: 2034082

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+n2p.model -DEV=xlip
-Q=/cgcn2.1/USPTO.spool_p/US09976782/runat_31082004_115951_10315/app_query.fasta_1.1351
-DB=SPTEMBL_25 -QFMT=faстан -SUFFIX=n2p.rspt -MINMATCH=0.1 -LOOPEXT=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=blomsum62 -TRANS=human40.cdi
-LIST=45 -DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09976782.@CGN_1_213 -runat_31082004_115951_10315 -NCFU=6 -ICPU=3
-NO MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSBLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

SPTEMBL_25.*
1: sp_archaea.*
2: sp_bacteria.*
3: sp_fungi.*
4: sp_human.*
5: sp_invertebrate.*
6: sp_mammal.*
7: sp_mhc.*
8: sp_organelle.*
9: sp_phase.*
10: sp_plant.*
11: sp_rodent.*
12: sp_virus.*
13: sp_vertebrate.*
14: sp_unclassified.*
15: sp_rvirus.*
16: sp_bacteriap.*
17: sp_archaeap.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1050.5	48.8	346	13 Q7ZZ32	Q7ZZ32 brachydanio

ID	Q7ZZ32	PRELIMINARY;	PRT;	346 AA.
AC	Q7ZZ32			
DT	01-JUN-2003	(TrEMBLrel. 24, Created)		
DT	01-JUN-2003	(TrEMBLrel. 24, Last sequence update)		
DT	01-OCT-2003	(TrEMBLrel. 25, Last annotation update)		
DE	SI:CI2P8.3	(Novel protein similar to human 5-hydroxytryptamine (Serotonin) receptor 5A (HTR5A)).		
GN	SI:CI2P8.3.			
OS	Brachydanio rerio (zebrafish) (Danio rerio).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;			
OC	Cyprinidae; Danio.			
OX	NCBI_TaxID=7955;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RA	Corby N.;			
RL	Submitted (FEB-2003) to the EMBL/GenBank/DBJ databases.			
DR	EMBL: AL772146; CAD61100.1; -			
DR	GO: GO:0016021; C:integral to membrane; IEA.			
DR	GO: GO:0001584; F:rhodopsin-like receptor activity; IEA.			
DR	GO: GO:0007186; P:G-protein coupled receptor protein signalin. .; IEA.			
DR	InterPro: IPR000276; GPCR_Rhodpsn.			
DR	Pfam: PF00001; 7tm_1; 1.			

Db	279	LeuIleGlyValPheValLeuCysTrpIleProPhePheLeuAlaGluLeuIlePro	298
Qy	998	CTCTGTGCCTGAGCTGCCCCCATCTGGAAAGCATATTTCTGTGGCTTGGCTACTCC	1057

Qy	1058	AAATCTTTCTTCAACCCCTGATTTACACAGCTTTTAAACAGACTACACATGCCTTC	1117
Db	319	ASnSerPhePheAsnProeuIeIyTrAlaPheAsnLyAsnTyraSnAsnAlaPhe	338
Qy	1118	AAGAGCCTCTTTACTAAGCAGAGA	1141

2025-01-01 10:00:00

Q86UNI	PRELIMINARY;	PRT; 247 AA.
ID	Q86UNI	
AC	Q86UNI	
DT	01-JUN-2003	(TREMELrel. 24, Created)
DT	01-JUN-2003	(TREMELrel. 24, Last sequence update)
DT	01-OCT-2003	(TREMELrel. 25, Last annotation update)
DE	Hypotheical protein HTR5A (Fragment).	

OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
NCBI_TaxID=9606;
[1]
SEQUENCE FROM N.A.
RP MEDLINE=99063792; PubMed=9847074;
RX

RT "Toward a complete human genome sequence." ;
RL Genome Res. 8:1097-1108(1998).
RN [2]
RP SEQUENCE FROM N.A.
RA Cotton M., Phillips A., Tomlinson C.;
RT "The sequence of Homo sapiens BAC clone RP11-5C23." ;
RL Submitted (SEP-2001) to the EMBL/GenBank/DBJ databases.

RP SEQUENCE FROM N. A.
RA Waterston R. H.;
PA Submitted (FEB-2002) to the EMBL/GenBank/DBJ databases.
[4]
RN
RP SEQUENCE FROM N. A.
RA Waterston R.;
PA Submitted (FEB-2002) to the EMBL/GenBank/DBJ databases.

RP SEQUENCE FROM N. A.
RA Waterston R.;
RL Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.
RN [6]
RP SEQUENCE FROM N. A.
RA Waterston R.;
RL Submitted (APR-2003) to the EMBL/GenBank/DBJ databases.

DR GO: GO:0016021; C:integral to membrane; IEA.
DR GO: GO:0001584; F:rhodopsin-like receptor activity; IEA.
DR GO: GO:0007100; F:transcription factor activity; IEA.
DR GO: GO:0007186; P:G-protein coupled receptor protein signalin...; IEA.
DR GO: GO:0006355; P:regulation of transcription, DNA-dependent; IEA.
DR InterPro: IPR000216; GPCR_Rhodopsn.
DR InterPro: IPR000847; HTH_LysR.

DR PRINTS; P00237; GPCRHDOPSN.
DR PROSITE; PS00237; G_PROTEIN_RECP_F1_1; 1.
DR PROSITE; PS0262; G_PROTEIN_RECP_F1_2; 1.
DR PROSITE; PS00044; HTH_LYSR_FAMILY; 1.
DR PROSITE; PS00044; HTH_LYSR_FAMILY; 1.
Hypothetical protein.
KW NON_TER 247 247
FT NON_TER 247 247
SQ SEQUENCE 247 AA; 38771F8B8F0137B6 CRC64;

Alignment Scores:	
Pred. No.:	4.82e-33
Length:	247

Score: 615.00 Matches: 143
 Percent Similarity: 67.35% Conservativity: 22
 Best Local Similarity: 58.37% Mismatches: 74
 Query Match: 28.55% Indels: 6
 DB: 4 Gaps: 4

US-09-976-782-15 (1-1152) x Q86UN1 (1-247)

QY 65 ACCAGCAGCGGACCGGACCCAGCCAGGAGGATACCTCGTTTCAGCCCGAGCGGC 124
 DB 8 ThrSerPheSerLeuSerThrProSerProLeuGluThrAsnHisSerLeuGlyLysAsp 27
 QY 125 GCCGTCCTGCGCGCGGAGGCGCCCTCTCTCTTCACGCTCCTGCTGTGTGAGCGTG 184
 DB 28 AspLeuArgProSer---SerProLeuLeuSerValPheGlyValLeuLeuLeuLeu 46
 QY 185 CTAGTGCTGCTGATCGCTCCACCTCTCTGTGGAACTGCTGTGTCGGTCCGCTCAGCATCCG 244
 DB 47 LeuGlyPheLeuValAlaAlaThrPheAlaTrpAsnLeuLeuValAlaThrIleLeu 66
 QY 245 CGGCTCCGCTGCTTCCACCGGCTGCGCATAACTTGGTGGCTCGACGCGCTCTCGGAC 304
 DB 67 ArgValArgThrPheHisArgValProHisnLeuValAlaSerMetAlaValSerAsp 86
 QY 305 GAATAGTGGCAGCGCTGGGATGACCCAGCGCTCGGAGTGAAGTCTGACACCGGCGGA 364
 DB 87 ValLeuValAlaAlaLeuValMetProLeuSerLeuValHisGluLeuSer---GlyArg 105
 QY 365 CGTGGCTGCTGGCGGAGCGCTGACAGTGGATCTCTTCGACGCGCGGAGCGCTGT 424
 DB 106 ArgTrpGlnLeuGlyArgLeuLeuGlnLeuTrpIleAlaCysAspVal----- 122
 QY 425 CTGTGCTGCCCCCGGCTCGGAACTGCGGCGCATCGCCCTGGCGCGGAGCGGCGCC 484
 DB 123 LeuCysCysThrAlaSerIleTrpAsnValThrAlaIleAlaLeuAspArgTrpSer 142
 QY 485 ATCACAGGACCTGACACACGCTGCGACCCGCGAGCGCGCTCTGCTCTCATGATC 544
 DB 143 IleThrArgHisMetGluTrpThrLeuArgThrArgLysCysValSerAsnValMetile 162
 QY 545 GCGTCCCGCGGTCGCTGCGGCTCATCCCTCGCGCGCTGCTCTTTGGCGGCGGC 604
 DB 163 AlaLeuThrTrpAlaLeuSerAlaValIleSerLeuAlaProLeuLeuPheGlyTrpGly 182
 QY 605 GAGGTGTGCGACGCTCGCTCCAGCGTCCAGGTGAGCGGGAACCTCTATGCGCGCC 664
 DB 183 GluThrTrpSerGluGlySerGluGluCysGlnValSerArgGluProSerTrpAlaVal 202
 QY 665 TTCTCCACCGCGCGCTTCCACCTGCGCTTGGCGTGGTGGTGGTGGTGGTGGTGGTGG 724
 DB 203 PheSerThrValGlyAlaPheTrpLeuProLeuLeuValLeuPheValTrpLys 222
 QY 725 ATCTACAGCGGCGCAAGTTTCTGTTTCGGC---CGCGCGGAGAGTGTGCTCCGTTG 781
 DB 223 IleTrpLysAlaAlaLysPheArgValGlySerArgLysThrAsnSerValSerProile 242
 QY 782 CCGGCCACCATGCGAG 796
 DB 243 SerGluAlaValGlu 247

RESULT 3

Q9N2B7 PRELIMINARY; PRT; 390 AA.
 AC Q9N2B7
 DT 01-OCT-2000 (TrEMBLrel. 15, Created)
 DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
 DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
 DE 5-hydroxytryptamine (Serotonin) receptor 1B.
 GN HTR1B.
 OS Gorilla gorilla (gorilla).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Gorilla.
 OC NCBI_TaxID=9593;

RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-gorilla-UI;
 RA Kitano T., Kobayakawa H., Saitou N.;
 RT "Silver Project";
 RL Submitted (APR-2000) to the EMBL/GenBank/DBJ databases.
 CC -!- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (BY SIMILARITY).
 CC -!- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
 DR EMBL; AB041372; BAA94457.1; -;
 DR GO; GO:0016021; C:integral to membrane; IEA.
 DR GO; GO:0004872; F:receptor activity; IEA.
 DR GO; GO:0001584; F:rhodopsin-like receptor activity; IEA.
 DR GO; GO:0007186; P:G-protein coupled receptor protein signalin. . .; IEA.
 DR InterPro; IP0000276; GPCR_Rhodopsn.
 DR Pfam; PF00001; 7tm_1; 1.
 DR PRINTS; PR00237; GPCRHHODOPSIN.
 DR PROSITE; PS00237; G-PROTEIN RECEPTOR FL1; 1.
 DR PROSITE; PS0262; G-PROTEIN RECEPTOR FL2; 1.
 KW G-protein coupled receptor; Receptor; Transmembrane.
 SQ SEQUENCE 390 AA; 43565 MW; 37D54B95E3277FC3 CRC64;

Alignment Scores:

Pred. No.: 4.45e-22 Length: 390
 Score: 451.00 Matches: 127
 Percent Similarity: 48.35% Conservativity: 63
 Best Local Similarity: 32.32% Mismatches: 147
 Query Match: 20.94% Indels: 56
 DB: 6 Gaps: 13

US-09-976-782-15 (1-1152) x Q9N2B7 (1-390)

QY 59 CCCGAGACCCAGCGGACCGGACCCAGCCAGCGGAGGATACCTCGTTTCGACCCGC 118
 DB 10 ProPro***ProAlaGlySerGluThrTrpValProGlnAlaAsnLeuSerSerAlaPro 29
 QY 119 AGC-----GGCGCGCTCTCCCGGCGGA 142
 DB 30 SerGlnAsnCysSerAlaLysAspTyrIleTyrGlnAspSerIleAlaLeuPro----- 47
 QY 143 GGGCGCGCTCTCTCTGCTTTCACGCTCCTGGTGTGAGCTGTGCTGTGCTGTGCTGTGCT 202
 DB 48 -----TrpLysValLeuLeuValMetLeuLeuAlaLeuIleThrLeu 61
 QY 203 GCACATTTCTGTGGAACCTGCTGCTTCCGGTTCACCATCCCGCGGTCGCTTCAC 262
 DB 62 AlaThrThrLeuSerAsnAlaPheValIleAlaThrValTyrArgThrArgLysLeuHis 81
 QY 263 CGCGTGGCGCATAACTTGGTGGCTCGACGCGCGCTCTCGGACGAACCTAGTGGAGCGGTG 322
 DB 82 ThrProAlaAsnTyrLeuIleAlaSerLeuAlaValThrAspLeuLeuValSerIleLeu 101
 QY 323 CGGATGCCAGCGGCTGGGAGTAGTGTGTGACCGGCGGAGCTGTGCTGTGCTGTGCTGTGCT 382
 DB 102 ValMetPro-----IleSerThrMetTyrThrValThrGlyArgTrpThrLeuGlyGln 119
 QY 383 AGCTGTGCGACGCTGTGATCTCTTTCGACGCGCGGAGCTGTGCTGTGCTGTGCTGTGCTGTG 442
 DB 120 ValValCysAspPheTrpLeuSerSerAsp-----IleThrCysCysThrAlaSer 136
 QY 443 CTGGGAACCTGGCGGCTCGCTCGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 502
 DB 137 IleLeuHisLeuLeuValIleAlaLeuAspArgTyrTrpAlaIleThrAspAlaValGlu 156
 QY 503 CACAGCTGGCAGCCCGGCGGCGGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTG 562
 DB 157 TyrSerAlaLysArgThrProLysArgAlaAlaValMetIleAlaLeuValTrpValPhe 176
 QY 563 TCGCGCTCATCGCTCGCGGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTG 622
 DB 177 SerIleSerIleSerLeuProPro---PhePheTrpArgGlnAlaLysAlaGluGluGlu 195
 QY 623 CTCAGCGCTGCGAGGTGAGCGGGAACCC---TCCTATCGCGCTTCTCCACCGCGGCGG 679

```

Db      196 ValSerGluCysValValAsnThrAspHisIleLeuTyrThrValTyrSerThrValGly 215
QY      680 GCCTTCCACCTGCGCGTGGCGTGTGTCTACCGGAAGATCTACAGCGCGCC 739
Db      216 AlaPheTyrPheProThrLeuLeuLeuIleAlaLeuTyrGlyArgIleTyrValGluAla 235
QY      740 AAGTTTCGTTTC-----GGCCGCGC-----CGGAGAGCTGTG 772
Db      236 ArgSerArgIleLeuLysGlnThrProAsnArgThrGlyLysArgLeuThrArgAlaGln 255
QY      773 CTG-----CCGTTGCGCGCCACCATGCGAGTGGAGTCCAAAGTAAAG 814
Db      256 LeuIleThrAspSerProGlySerThrSerValThrSerIleAsnSerArgValPro 275
QY      815 GAAGCACCTGATGAGCT---GAAGTGGTGTTCAGCGGCACATGTCAAAGCAACGCTGTCC 871
Db      276 AspValProSerGluSerGlySerProValTyrValAsnGlnValLysValArgValSer 295
QY      872 TTCAGGTGAGCGGGACTCTCGCGGAGCAGAGAAG-----GAGAGG 913
Db      296 -----AspAlaLeuLeuGluLysLysLysLeuMetAlaAlaArgGluArg 310
QY      914 CGAGCAGCATGATGTGGGAATCTGATGGCGTGTGTCTGTGTGTGTGTGTGTGTGTGTGTGT 973
Db      311 LysAlaThrLysThrLeuGlyIleIleLeuGlyAlaPheIleValCysTrpLeuProPhe 330
QY      974 TTCCTGACGAACACTATCAGCCACTCTGT-----GCTGAGCGCTGCCCCCATCTGG 1027
Db      331 PheIleIleSerLeuValMetProIleCysLysAspAlaCysTrpPheHisLeuAlaIle 350
QY      1028 AAAAGCATATTTCTGGTGTGGTCTACTCCAAATTTCTTCAACCCCTGATTTACACA 1087
Db      351 PheAspPhePheThrTrpLeuGlyTyrLeuAsnSerLeuLeuAsnProIleIleTyrThr 370
QY      1088 GCTTTTAAACAAGAACTACAACTGCTTCAAGAGCCTC 1126
Db      371 MetSerAsnGluAspPheLysGlnAlaPheHisLysLeu 383

RESULT 4
Q99MB0 PRELIMINARY; PRT; 386 AA.
AC Q99MB0;
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Serotonin receptor 5HT1B.
OS Mesocricetus auratus (Golden hamster).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Cricetinae;
OC Mesocricetus.
OX NCBI_TaxID=10036;
RN [1]
RP SEQUENCE FROM N.A.
RA Devor E.J., Devor R.M.;
RT "Cloning and sequencing the serotonin receptor 5HT1B in the Golden
RT Hamster.";
RL Submitted (MAR-2001) to the EMBL/GenBank/DBJ databases.
CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (BY SIMILARITY).
CC -1- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
DR EMBL; AF356554; AAK25827.1; -.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0004872; F:receptor activity; IEA.
DR GO; GO:0001584; F:rhodopsin-like receptor activity; IEA.
DR GO; GO:0007186; P:G-protein coupled receptor protein signalin. .; IEA.
DR InterPro; IPR000276; GPCR_Rhodpsn.
DR Pfam; PF00001; 7tm 1; 1.
DR PRINTS; PR00237; GPCR_RHODOPSIN.
DR PROSITE; PS00217; G PROTEIN RECEPTOR FL 1; 1.
DR PROSITE; PS00262; G-PROTEIN RECEPTOR FL 2; 1.
KW G-protein coupled receptor; Receptor; Transmembrane.
SQ SEQUENCE 386 AA; 43020 MW; 19F430C36F5CEB40 CRC64;

```

Alignment Scores:

```

Pred. No.: 3,29e-21 Length: 386
Score: 438.00 Matches: 120
Percent Similarity: 51.30% Conservative: 57
Best local Similarity: 34.78% Mismatches: 132
Query Match: 20.33% Indels: 36
DB: 11 Gaps: 12

US-09-976-782-15 (1-1152) x Q99MB0 (1-386)

QY      167 GTCCTGGTGTGACGCTGCTAGTGTGCTGATCGTCCACTTTCTCTGTGGAACTGCTG 226
Db      46 ValLeuLeuValAlaLeuLeuAlaLeuIleThrLeuAlaThrThrLeuSerAsnAlaPhe 65
QY      227 GTTCCGGTCACCATCCCGCGTCCGTGCTCTTCCACCGCGTCCCGCATATACTTGTGGCC 286
Db      66 ValIleAlaThrValTyrArgThrArgLysLeuHisThrProAlaAsnTyrLeuIleAla 85
QY      287 TCGACGGCGCTTCGAGCAACTAGTGGCAGCGCTGGGATGCCACCGACCTGCGGAGT 346
Db      86 SerLeuAlaValThrAspLeuValSerIleLeuValMetProValSerThrMetTyr 105
QY      347 GAGCTGTCCACGGCGGACGCTCGGCTGCTGGCGGAGCGCTGTGCCACGCTGTGGATCTCC 406
Db      106 ThrVal---ThrGly---ArgTrpThrLeuGlyGlnValValCysAspPheTrpLeuSer 123
QY      407 TTCGACGGCGGAGCGCTGCTGTGCTGCCCCCGCGCGCTCGGAAACGTGGCGGCATTCGCC 466
Db      124 SerAsp-----IleThrCysCysThrAlaSerIleMetHisLeuCysValIleAla 140
QY      467 CTGGCGCGCGACGGGCGCATCACAGGCACCTGTCGACACACGCTGGCGGACCGCGCCG 526
Db      141 LeuAspArgTyrTrpAlaIleThrAspAlaValGluTyrSerAlaLysArgThrProLys 160
QY      527 GCTCGTGTCTCATGATCGCTCGCGCGGCGCTGCGCGCTCATCGCCCTCGCGCGCG 586
Db      161 ArgAlaAlaIleMetIleAlaLeuValTrpValPheSerIleSerLeuLeuProPro 180
QY      587 CTGCTCTTTGGCGCGGAGGTGTGCGACGCTCGCGCTCCAGCGTCCAGCGTGGCAGTGGACCGG 646
Db      181 ---PhePheTrpArgGlnAlaLysAlaGluGluValLeuValLeuAspCysPheValAsnThr 199
QY      647 GAA---CCCTCTATGCGCCCTTCCACCGCGCGCGCTTCCACCTGCGCGTGGCGGTG 703
Db      200 AspHisValLeuTyrThrValTyrSerThrGlyAlaPheTyrLeuLeuProThrLeuLeu 219
QY      704 GTGCGGTTGTCTACCGGAAGATCTACGAGCGCGCCAAAGTTTCGTTTC-----751
Db      220 LeuIleAlaLeuTyrGlyArgIleTyrValGluAlaArgSerArgIleLeuLysGlnThr 239
QY      752 -----GGCGCGCG---CGGAGAGCTGTGCTG-----CCG 778
Db      240 ProAsnLysThrGlyLysArgLeuThrArgAlaGlnLeuIleThrAspSerProGlySer 259
QY      779 TTGCGCGCGCACCATGTCAGTGGTCCAGGTCCAAAGTAAGAACGACCTGATGAGGCT---GAA 835
Db      260 ThrThrSerValThrSerIleAsnSerArgAlaProAspValProSerGluSerGlySer 279
QY      836 GTGCTGTTCAGGCACATTTGCAAGCAACGCTGCTCTTCCAGGTGAGCGGGGACTCTCTGG 895
Db      280 ProValTyrValAsnGlnValLysValArgValSer-----AspAlaLeu 294
QY      896 CGGAGCAGAGAAG-----GAGAGCGAGCAGCCATGATGTGGGAATT 937
Db      295 LeuGluLysLysLysLeuMetAlaAlaArgGluArgLysAlaThrLysThrLeuGlyIle 314
QY      938 CTGATTTGGCTGTTGTGCTGTGATCCCTTCTTCTCTGACGGAACCTCATCAGGCCCA 997
Db      315 IleLeuGlyAlaPheIleValCysTrpLeuProPhePheIleIleSerLeuValMetPro 334
QY      998 CTCTGT-----GCCTGCGAGCCTGCCCGCCCATCTGGAAGAGCATATTCTGTGCTTGGC 1051
Db      335 IleCysLysAspAlaCysTrpPheHisMetAlaIlePheAspPheAsnTrpLeuGly 354

```


OS Aedes aegypti (Yellowfever mosquito).
 OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
 OC Neoptera; Endopterygota; Diptera; Nematocera; Culicoidea; Aedes.
 OX NCBI_TaxID=71159;
 RN [1]
 RP SEQUENCE FROM N.A.
 RT Pietrantoni P.V., Jagge C.L.;
 RT "Cloning of a serotonin receptor 5HT7-like cDNA from mosquito Aedes
 RT aegypti (Diptera: Culicidae) excretory and respiratory systems.";
 RL Submitted (JAN-2001) to the EMBL/GenBank/DBS databases.
 CC -!- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (BY SIMILARITY).
 CC -!- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
 DR EMBL; AF296125; AAG49292.1; -
 DR GO; GO:0016021; C:integral to membrane; IEA.
 DR GO; GO:0004872; F:receptor activity; IEA.
 DR GO; GO:0001584; F:rhodopsin-like receptor activity; IEA.
 DR GO; GO:0007186; P:G-protein coupled receptor protein signalin. . .; IEA.
 DR InterPro; IPR000276; GPCR_Rhodopsn.
 DR Pfam; PF00001; 7tm 1; 1.
 DR PRINTS; PR00237; GPCRRHODOPSN.
 DR PROSITE; PS00237; G-PROTEIN RECEPTOR FL_1; 1.
 DR PROSITE; PS00262; G-PROTEIN RECEPTOR FL_2; 1.
 KW G-protein coupled receptor; Receptor; Transmembrane.
 SQ SEQUENCE 464 AA; 50821 MW; 3F432DA22E0345E4 CRC64;
 Alignment Scores:
 Pred. No.: 5 69e-21 Length: 464
 Score: 434.50 Matches: 118
 Percent Similarity: 49.35% Conservative: 71
 Best Local Similarity: 30.81% Mismatches: 163
 Query Match: 20.17% Indels: 31
 DB: 5 Gaps: 9
 US-09-976-782-15 (1-1152) x Q9GQ54 (1-464)
 QY 11 GCCGCTAGCTTTAGTGGCCACCGCGCGTTCCTTCCCTTGCCTTGCCTGGGA----- 58
 Db 58 SerSerSerAlaSerValSerAlaThrAlaIleAlaThrSerLeuProAlaLeuValAsp 77
 QY 59 -----CCGAGACACGAGCGGACCGCGGACCCCGAGAGGAGGATCTCGGTTG 112
 Db 78 ArgLeuThrProThrSerSer-----ThrSerSerLeuLeuAspGluLeuGly 94
 QY 113 ACCCGAGCGCGCGTCTCGCGGCGCGGCGCGCTTCTCTGTCTTACGGTCCG 172
 Db 95 ThrSerGluSerSer-----ProAlaGluProValAsnValLeuThrIleGln 110
 QY 173 GTGGTACCGTGTAGTGTGTGTGATCGCT-----GCCACTTCTGTGGAACTCG 223
 Db 111 ThrIleValIleSerIleValLeuLeuAlaValIleIleGlyThrIleValGlyAsnVal 130
 QY 224 CTGTTCCGCTACCATCCGCGGTCGCGTCCACCGCGTCCCGGATACCTTGGTG 283
 Db 131 LeuValCysValAlaValCysLeuValArgLysLeuArgArgProCysAsnTyrLeuLeu 150
 QY 284 GCCTCGACGCGCGTCTCGACGAACTAGTGGAGCGCTGGCGATGCCACCGCGCGG 343
 Db 151 ValSerLeuAlaIleSerAspLeuValAlaValValMetProProAlaLeuLeu 170
 QY 344 AGTGAAGCTTCGACCGCGGCGAGTGTGCTGTGGCGGAGCGCTGTGCCACGCTGTGATC 403
 Db 171 TyrGluValLeuGluGluTrpLys-----PheGlyThrValPheCysAspIleTrpVal 188
 QY 404 TCTTCGACCGCGGACCTGTCTGTCTGCTCCCGCGCGCTCGGAACTGGCGGCGATC 463
 Db 189 SerPheAspVal-----LeuSerCysThrAlaSerIleLeuAsnLeuCysAlaIle 205
 QY 464 GCCTGGCGCGGACCGCGGCGCATCACACGCACTGCAGCACACAGCTGGCGCACCCGACG 523
 Db 206 SerValAspArgTyrTrpAlaIleThrLysProLeuGluTyrGlyValLysArgThrPro 225
 QY 524 CCGCGCTGCTGTATGATGATGCGCTGCGCGCGGCTGCGCTGCGCTGCTGCGCTGCGG 583
 Db 524 CCGCGCTGCTGTATGATGATGCGCTGCGCGCGGCTGCGCTGCGCTGCTGCGCTGCGG 583

Db 226 ArgArgMetIleAlaCysIleValLeuValTrpLeuValAlaAlaCysIleSerLeuPro 245
 QY 584 CGCGTGTCTTTGGCGGGCGAGGTGTGCGACGCTCGGCTCCAGCGCTCCAGGTGAGC 643
 Db 246 ProLeuLeuIleLeuGlyAsnGluHisMetThrAsnGlyGlnProSerCysSerValCys 265
 QY 644 CGGGAACCTCTCTATCGCGCTTCTCCACCGCGCGCTTCCACCTGCGCTGCGGTG 703
 Db 266 GlnAsnPhePheTyrGlnIleTyrAlaThrLeuCysAlaPheTyrIleProLeuAlaVal 285
 QY 704 GTGCGCTTTGTCTACCGAAGATCTCAGGCGGCGCAAGTTTCGTTTCGCGCGCGCG 763
 Db 286 MetLeuPheValTyrPheGlnIlePheArgAlaAlaArgArgIleValAsnGluGluLys 305
 QY 764 AGAGCTGTGCTGCTGCGCGTTCGCGCGCACCATGCGAGGTGAGGTCCAGGTAAAGAACACCT 823
 Db 306 ArgAlaGlnLysHisLeuGluThrAlaIleAsnGlySerAlaThrThrProGluLysLys 325
 QY 824 GATGAGCT-----GAAGTGTGTTCACGGCACCATGCAAGCAACCGTGTCTTCCAG 877
 Db 326 LeuSerAlaGlyGlyThrValLeuValAlaThrProGlnHisLysArgLeuArgPheGln 345
 QY 878 GTGACGCGGACTCTCTGCGGAGCAGAGGAGCGGAGCGGAGCCATGATGTGGGAATT 937
 Db 346 LeuAla-----LysGluArgLysAlaSerThrThrLeuGlyIle 358
 QY 938 CTGATTGCGGTGTTGTGTGTGCTGCTGCGTCCCTTCTTCCACGGAACATCATCAGCCA 997
 Db 359 IleMetSerAlaPheThrValCysTrpLeuProPheIleLeuAlaLeuValArgPro 378
 QY 998 CTCTGTGCTGCGAGCTGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1057
 Db 379 PheLeuGly---GluAspHisHisLeuLeuSerSerLeuPheLeuTrpLeuGlyTyrAla 397
 QY 1058 AATCTTCTTCAACCCCTGATTTACACAGCTTTTACAGAACTACACAACTGCTTCTC 1117
 Db 398 AsnSerLeuLeuAsnProIleIleTyrAlaThrLeuAsnArgAspPheArgLysProPhe 417
 QY 1118 AAGAGCCTC 1126
 Db 418 GlnGluIle 420
 RESULT 7
 O76267 PRELIMINARY; PRT; 492 AA.
 AC O76267;
 DT 01-NOV-1998 (TREMBLrel. 08, Created)
 DT 01-NOV-1998 (TREMBLrel. 08, Last sequence update)
 DT 01-JUN-2003 (TREMBLrel. 24, Last annotation update)
 DE G-protein-coupled 5-hydroxytryptamine receptor.
 GN 5-HTA1.
 OS Aplysia californica (California sea hare).
 OC Eukaryota; Metazoa; Mollusca; Gastropoda; Orthogastropoda;
 OC Apogastropoda; Heterobranchia; Euthyneura; Opisthobranchia; Anaspidia;
 OC Aplysioidae; Aplysiidae; Aplysia.
 OX NCBI_TaxID=6500;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=98337861; PubMed=9671650;
 RA Angers A., Storzuk M.V., Duchaine T., Castellucci V.F.,
 RA Desgosses L.,
 RT "Cloning and functional adenylation of an aplysia 5-HT receptor
 RT negatively coupled to adenylylate cyclase.";
 RL J. Neurosci. 18:5586-5593(1998).
 CC -!- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (BY SIMILARITY).
 CC -!- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
 DR EMBL; AF041039; AAC28786.1; -
 DR GO; GO:0016021; C:integral to membrane; IEA.
 DR GO; GO:0004872; F:receptor activity; IEA.
 DR GO; GO:0001584; F:rhodopsin-like receptor activity; IEA.
 DR GO; GO:0007186; P:G-protein coupled receptor protein signalin. . .; IEA.
 DR InterPro; IPR000276; GPCR_Rhodopsn.
 DR Pfam; PF00001; 7tm 1; 1.

US-09-976-782-15 (1-1152) x Q9N297 (1-422)

QY 134 CGGGCCGAGGG-----CGCCCTTCTCT----- 157
Db 6 ProGlyGlnGlyAsnAsnThrThrSerProProAlaProPheGluThrGlyGlyAsnThr 25
QY 158 -----GTCATCAGCTGCTGGTGTGACGCTGCTAGTGTG----- 193
Db 26 ThrGlyIleSerAspValThrPheSerTyrGlnValIleThrSerLeuLeuLeuGlyThr 45
QY 194 CTGATCGTGCACCTTCTGCTGGAACTGCTGCTGGTCCGGTACCATCCCGGGTCCGT 253
Db 46 LeuIlePheCysAlaValLeuGlyAsnAlaCysValValAlaAlaIleAlaLeuGluArg 65
QY 254 GCCTTCCACCGGCTCGGCACTTGTGGCTCGACGCGCTCGCGACCACTAGTG 313
Db 66 SerLeuGlnAsnValAlaAsnTyrLeuIleGlySerLeuAlaValThrAspLeuMetVal 85
QY 314 GCAGCGCTGGCGATGCCAGCGCTGGCGAGTGTGTCGACCGGGGAGCTCGGCTG 373
Db 86 SerValLeuValLeuPro-----MetAlaAlaLeuTyrGlnValLeuAsnLysTrpThr 103
QY 374 CTGGGCGGAGCTGTGCGACGTGTGATCTCTTCGACCGCGGACCTGTCTGTCTGC 433
Db 104 LeuGlyGlnValThrCysAspLeuPheIleAlaLeuAspVal-----LeuCysCys 120
QY 434 CGCGCGGCTCGGGAACGTGGCGCATCGCGCTGGCGCGGCGGCATCACACGG 493
Db 121 ThrSerSerIleLeuHisLeuCysAlaIleAlaLeuAspArgTyrTrpAlaIleThrAsp 140
QY 494 CACCTCGACACACGCTGGCGACCCCGACCGCGCTGCTGTGCTATGATCGCGCTCGC 553
Db 141 ProIleAspTyrValAsnLysArgThrProArgAlaAlaAlaLeuIleSerLeuThr 160
QY 554 CGGGTCCGCTCGCGCTGCTCGCCCTCGCGCGCTGCTTGGCGGGCGGAGTGTGC 613
Db 161 TrpLeuIleGlyPheLeuIleSerIleProProMetLeuGlyTrpArgThrProGluAsp 180
QY 614 GACGCTCGGCTCGACGCTGCCAGTGCAGCGGGAACCTCTATCGCGCTTCTCCAC 673
Db 181 ArgSerAspProAspAlaCysThrIleSerLysAspHisGlyTyrThrIleTyrSerThr 200
QY 674 CGGGGCGCTTACCTGCGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCT 733
Db 201 PheGlyAlaPheTyrIleProLeuLeuLeuMetLeuValLeuTyrGlyArgIlePheArg 220
QY 734 GCGGCCAAGTTCTGTTTCGCGCG-----CGC 760
Db 221 AlaAlaArgPheArgIleArgLysThrValLysValGluLysThrGlyAlaAspThr 240
QY 761 CGGAGAGCTGTGCTGCGCTGCGCGCG----- 787
Db 241 ArgHisGlyAlaSerProAlaProGlnProLysLysSerValAsnGlyGluSerGlySer 260
QY 788 -----ACATCGAGTGTAGGTCCAAG----- 808
Db 261 ArgAsnTrpArgLeuGlyValGluSerLysAlaGlyGlyAlaLeuCysAlaAsnGlyAla 280
QY 809 GTAAGGAAGACCTCATGAGGCTGAAGTG---GTGTTCCAGCGCACAT----- 853
Db 281 ValArgGlnGlyAspAspGlyAlaAlaLeuGluValIleGluValHisArgValGlyAsn 300
QY 854 -----TGCAAGACCAACGGTG 868
Db 301 SerLysGluHisLeuProLeuProSerGluAlaGlyProThrProCys---AlaProAla 319
QY 869 TCTTCCAGTGTAGCGGACTCTCTGGCGGACCAAG----- 907
Db 320 SerPheGluArgLysAsnGluArgAsnAlaGluAlaLysArgLysMetAlaLeuAlaArg 339
QY 908 GAGAGCGGAGCAGCCATGATGGTGGGAATTCATTGGCGTGTCTTGTCTGTCTGGATC 967
Db 340 GluArgLysThrValLysThrLeuGlyIleMetGlyThrPheIleLeuCysTrpLeu 359

QY 968 CCCTTCTCTCTGACGGAACATCATCAGCCCACTCTGT-----GCCTGCAAGCTGCCCCC 1021
Db 360 ProPhePheIleValAlaLeuValLeuProPheCysGluSerSerCysHisMetProThr 379
QY 1022 ATCTGGAAGAAGCATATTTCTGTGGCTGTGCTACTCAATCTTCTTCAACCCCTGATT 1081
Db 380 LeuLeuGlyAlaIleIleAsnTrpLeuGlyTyrSerAsnSerLeuLeuAsnProValIle 399
QY 1082 TACACAGCTTTTAAACAAGAACTACACAATGCCTTCAAGAGCTC 1126
Db 400 TyrAlaTyrPheAsnLysAspPheGlnAsnAlaPheLysLysIle 414
RESULT 9
O18512 PRELIMINARY; PRT; 445 AA.
AC O18512; Q9BKW9;
DT 01-JAN-1998 (TrEMBLrel. 05, Created)
DT 01-JAN-1998 (TrEMBLrel. 05, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Serotonin receptor (Hypothetical protein).
GN Y22D7AR.13.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC Rhabditidae; Peloderinae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN [1]
RP SEQUENCE FROM N.A.
RA Olde B., McCombie R.W.;
RT Molecular cloning and functional expression of a serotonin receptor
RT from Caenorhabditis elegans.;
RL J. Mol. Neurosci. 0:0-0(1997).
RN [2]
RP SEQUENCE FROM N.A.
RA Dubnick M.;
RL Submitted (SEP-1994) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=Bristol N2;
RX MEDLINE=99069613; PubMed=9851916;
RA None;
RT "Genome sequence of the nematode C. elegans: a platform for
RT investigating biology. The C. elegans Sequencing Consortium.";
RL Science 282:2012-2018(1998).
RN [4]
RP SEQUENCE FROM N.A.
RC STRAIN=Bristol N2;
RA Madsen C.;
RT "The sequence of C. elegans cosmid Y22D7AR.";
RL Submitted (OCT-2000) to the EMBL/GenBank/DBJ databases.
RN [5]
RP SEQUENCE FROM N.A.
RC STRAIN=Bristol N2;
RA Waterston R.;
RT "Direct Submission.";
RL Submitted (SEP-2001) to the EMBL/GenBank/DBJ databases.
CC -!- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (BY SIMILARITY).
CC -!- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
DR EMBL; U15167; AAB66360.1; -.
DR EMBL; AC084154; AAK29875.2; -.
DR FIR; T42203; T42203.
DR WormPep; Y22D7AR.13; CE29370.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0004872; F:receptor activity; IEA.
DR GO; GO:0001584; F:rhodopsin-like receptor activity; IEA.
DR GO; GO:0007186; P:G-protein coupled receptor protein signalin...; IEA.
DR InterPro; IPR000276; GPCR_Rhodpsn.
DR Pfam; PF00001; 7tm1; 1.
DR PRINTS; PR00237; GPCR_RHODPSN.
DR PROSITE; PS00237; G-PROTEIN RECEPTOR FL1; 1.
DR PROSITE; PS00262; G-PROTEIN RECEPTOR FL2; 1.
KW Hypothetical protein; G-protein coupled receptor; Receptor;
KW Transmembrane.


```
QY 374 CTGGCCGGAGCTGTGCCACGTGTGGATCTCTTCGACGGCGAGCTGTCTGTGCTGC 433
Db |||||... ||| : : : : : ||| |||||
QY 104 LeuGlyGlnValThrCysAspLeuPheIleAlaLeuAspVal-----LeuCysCys 120
Db : : : : : |||||
QY 434 CCCGCCGCTCGGAACGTTGGCGCCATTCGCCCTTGGCGCCGAGCGGCATCACGG 493
Db : : : : : |||||
QY 121 ThrSerSerIleLeuHisCysAlaIleAlaLeuAspArgTyrTrpAlaIleThrAsp 140
Db : : : : : |||||
QY 494 CACCTGACGACACGCTCGGACCCGCGAGCGCGCTGTGCTCATGATCGGCTGCC 553
Db : : : : : |||||
QY 141 ProfileAspTyrValAsnLysArgThrProArgArgAlaAlaLeuIleSerLeuThr 160
QY 554 CGGTGGCGCTCGGCTCATCGCCCTCGCGCGCTGCTCTTGGCGCGGCGAGGTGTC 613
Db : : : : : |||||
QY 161 TrpLeuIleGlyPheLeuIleSerIleProMetLeuGlyTrpArgThrProGluAsp 180
QY 614 GACCTCGGCTCCAGCGCTCGGAGTGGCGCGGAAACCTCTATGCGCGCTTCGCCAC 673
Db : : : : : |||||
QY 181 ArgSerAspProAspAlaCysThrIleSerLysAspHisGlyTyrThrIleTyrSerThr 200
QY 674 CGCGCGCTTCCACCTCGGCTTGGCGTGGCGTGGCTTGTCTACCGGAAGATCTACAG 733
Db : : : : : |||||
QY 201 PheGlyAlaPheTyrIleProLeuLeuMetLeuValLeuTyrGlyArgIlePheArg 220
QY 734 CGGCGCAAGTTTCGTTTCGGCGCGCGCGAGAGCTGTG----- 772
Db |||||... |||||
QY 221 AlaAlaArgPheArgIleArgLysThrValLysLysValGluLysThrGlyAlaAspThr 240
QY 773 -----CTCCGCTTGGCGGC----- 787
QY 241 HisHisGlyAlaSerProAlaProGlnProLysLysSerValAsnGlyGluSerGlySer 260
QY 788 -----ACATGCAGTGAGTCCAAAG----- 808
QY 261 ArgAsnTrpArgLeuGlyValGluSerLysAlaGlyGlyLeuCysAlaAsnGlyAla 280
QY 809 GTAAAGGAAGCACCTGTAGTGGCTGAAGTG---GTGTTACGGGCACAT----- 853
QY 281 ValArgGlnGlyAspAspGlyAlaAlaLeuGluValIleGluValHisArgValGlyAsn 300
QY 854 -----TGCAAAAGCAACGGTG 868
QY 301 SerLysGluHisLeuProLeuProSerGluAlaGlyProThrProCys---AlaProAla 319
QY 869 TCCTTCCAGTGAGCGGGACTCTCGCGGAGCAGAG----- 907
QY 320 SerPheGluArgLysAsnGluArgAsnAlaGluAlaLysArgLysMetAlaLeuAlaArg 339
QY 908 GAGAGCGGAGCAGCACCATGATGTGGGAATTCGTGTCGTGTTGTGCTGTGTCGATC 967
QY 340 GluArgLysThrValLysThrLeuGlyIleIleMetGlyThrPheIleLeuCysTrpLeu 359
QY 968 CCCTTCTTCGAGGAACATCATCAGCCCACTCTGT-----GCTGAGCGCTGCCCCC 1021
QY 360 ProphePheIleValAlaLeuValLeuProPheCysGluSerSerCysHisMetProThr 379
QY 1022 ATCTGGAAGAACATATTTCTGGCTTGGCTTACTTCCATTTCTTCAACCCCTGATT 1081
QY 380 LeuLeuGlyAlaIleIleAsnTrpLeuGlyTyrSerAsnSerLeuLeuAsnProValIle 399
QY 1082 TACACAGCTTTTAAACAAGCACTACAACATGCTTCAAGAGCGCTC 1126
QY 400 TyrAlaTyrPheAsnLysAspPheGlnAsnAlaPheLysLysIle 414
Db |||||... |||||
RESULT 12
Q9N298
ID Q9N298 PRELIMINARY; PRT; 422 AA.
AC Q9N298
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Serotonin receptor 1A.
GN HTR1A.
```

```
OS Pan troglodytes (Chimpanzee).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Pan.
OX NCBI_taxid=9598;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=chimp-220;
RA Kitano T., Kobayakawa H., Saitou N.;
RT "Silver Project.";
RL Submitted (APR-2000) to the EMBL/GenBank/DBJ databases.
CC -!- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (BY SIMILARITY).
CC -!- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
DR EMBL; AB041404; BAA94489.1; -.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0004872; F:receptor activity; IEA.
DR GO; GO:0001584; P:rhodopsin-like receptor activity; IEA.
DR GO; GO:0007186; P:G-protein coupled receptor protein signalin. .; IEA.
DR InterPro; IPR000276; GPCR_Rhodpsn.
DR Pfam; PF00001; 7tm.1; 1.
DR PRINTS; PR00237; GPCRHHODPSN.
DR PROSITE; PS00237; G_PROTEIN_RECEP_F1_1; 1.
DR PROSITE; PS0262; G_PROTEIN_RECEP_F1_2; 1.
KW G-protein coupled receptor; Receptor; Transmembrane.
SQ SEQUENCE 422 AA; 46171 MW; FCD989BD0313A1A0 CRC64;

Alignment Scores:
Pred. No.: 1,54e-20 Length: 422
Score: 428.00 Matches: 120
Percent Similarity: 45.22% Conservative: 69
Best Local Similarity: 28.71% Mismatches: 133
Query Match: 19.87% Indels: 96
DB: 6 Gaps: 13

US-09-976-782-15 (1-1152) x Q9N298 (1-422)
QY 134 CCGGCCCGAGGG-----CGCCCTTCTCT------ 157
Db |||||... |||||
QY 6 ProGlyGlnGlyAsnAsnThrThrSerProAlaProPheGluThrGlyGlyAsnThr 25
QY 158 -----GTCTTACGGTCTCTGGTGTGAGCGCTCTAGTGTG----- 193
Db |||||... |||||
QY 26 SerGlyIleSerAspValThrPheSerTyrGlnValIleThrSerLeuLeuLeuGlyThr 45
QY 194 CTGATCGCTGCCACTTTCTCTGTGGAACCTGTGCTCGGTCAACCATCCGCGGTCTCGT 253
Db |||||... |||||
QY 46 LeuIlePheCysAlaValLeuGlyAsnAlaCysValValAlaAlaIleAlaLeuGluArg 65
QY 254 GCCTTCCACCGCTGCCGATAACTTGTGTGGCTCTGACGGCGCTCTCGAGCAACTAGTG 313
Db |||||... |||||
QY 66 SerLeuGlnAsnValAlaAsnTyrLeuIleGlySerLeuAlaValThrAspLeuMetVal 85
QY 314 GCAGGCTGGCGATGCCACCGAGCTGGCGAGTGTGTGTGACCGCGGAGCGTGTGCGGTG 373
Db |||||... |||||
QY 86 SerValLeuValLeuPro-----MetAlaAlaLeuTyrGlnValLeuAsnLysTrpThr 103
QY 374 CTGGCCGAGCGCTGTGCCACGTGTGGATCTCTTCGACGGCGGAGCGCTGTGTGCTGC 433
Db |||||... |||||
QY 104 LeuGlyGlnValThrCysAspPheIleAlaLeuAspVal-----LeuCysCys 120
QY 434 CCGCGCGCTTCGGGACGTGGCGCCATTCGCCCTTGGCGCGGAGCGGCATCACAGG 493
Db |||||... |||||
QY 121 ThrSerSerIleLeuHisCysAlaIleAlaLeuAspArgTyrTrpAlaIleThrAsp 140
QY 494 CACCTGACGACACGCTGGCGGACCCGCGAGCGCGCTGTGCTCATGATCGGCTGCC 553
Db |||||... |||||
QY 141 ProfileAspTyrValAsnLysArgThrProArgArgAlaAlaLeuIleSerLeuThr 160
QY 554 CGGTGGCGCTCGGCTCATCGCCCTCGCGCGCTGCTCTTGGCGCGGCGAGGTGTC 613
Db |||||... |||||
QY 161 TrpLeuIleGlyPheLeuIleSerIleProMetLeuGlyTrpArgThrProGluAsp 180
QY 614 GACGCTCGGCTCCAGCGCTGAGTGGCGGGAACCTCTCTATGCGCGCTTCGCCAC 673
Db |||||... |||||
```


[illegible]

Search completed: August 31, 2004, 20:21:30
Job time : 186 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - protein search, using frame_plus_n2p model

Run on: August 31, 2004, 20:07:32 ; Search time 50.5 Seconds

(without alignments)

4388.619 Million cell updates/sec

Title: US-09-976-782-15

Perfect score: 2154

Sequence: 1 cgcctggagccgtagcc.....aagcagagatgaacacaggg 1152

Scoring table:

BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 283366 seqs, 96191526 residues

Total number of hits satisfying chosen parameters: 566732

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+ n2p.model -DEV=xlp
-Q/cgn2 1/USPTO spoel p/US09976782/runat 31082004.115951.10328/app.query.fasta_1.1351
-DB=pir 78 -QWMT=fastaan -SUFFIX=n2p.rpr -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=-1 -WATIRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USBR=US09976782 @CGN 1.1 64 @runat 31082004.115951.10328 -NCPU=6 -ICPU=3
-NO MAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSBLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREAS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

pir 78:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1452.5	67.4	370	2	I48231
2	1443	67.0	369	2	S38744
3	1063.5	49.4	357	2	I37107
4	1020.5	47.4	357	2	S28058
5	1014.5	47.1	357	2	B47472
6	449	20.8	389	2	S68422
7	448	20.8	390	2	JN0268
8	444	20.6	390	2	S58126
9	441	20.5	377	2	B30341
10	440	20.4	386	2	S54153
11	438	20.3	386	2	A42688
12	437	20.3	386	2	S18637
13	435.5	20.2	564	2	A38271
14	431.5	20.0	445	2	T42203

15	430	20.0	377	2	A53279	serotonin receptor
16	428	19.9	422	2	I38209	serotonin receptor
17	425	19.7	421	2	I49375	serotonin receptor
18	421	19.5	422	2	JH0315	serotonin receptor
19	417	19.4	445	2	A48881	serotonin receptor
20	416.5	19.3	509	2	A47174	serotonin receptor
21	416	19.3	374	2	I77467	serotonin receptor
22	416	19.3	377	2	S68423	serotonin receptor
23	412.5	19.2	366	2	A47321	serotonin receptor
24	411.5	19.1	366	2	S26048	serotonin receptor
25	409.5	19.0	379	2	JC6178	serotonin receptor
26	408	18.9	365	2	A45260	serotonin receptor
27	408	18.9	448	2	A47519	serotonin receptor
28	408	18.9	484	2	S58868	serotonin receptor
29	406.5	18.9	366	2	A47385	G protein-coupled
30	403	18.7	448	2	S36402	serotonin receptor
31	385.5	17.9	450	2	A38316	serotonin receptor
32	379	17.6	517	2	A45121	alpha-2-adrenergic
33	378	17.5	515	2	A40491	alpha-1B adrenergic
34	375.5	17.4	461	2	A31237	alpha-1-adrenergic
35	374	17.4	443	1	DYH022	alpha-2C-adrenergic
36	374	17.4	458	2	A40392	dopamine receptor
37	370.5	17.2	453	2	A35642	alpha-2B-adrenergic
38	370.5	17.2	477	2	S71323	alpha-2B-adrenergic
39	370	17.2	444	1	DYB002	alpha-1A adrenergic
40	369.5	17.2	400	2	G01977	dopamine receptor
41	369.5	17.2	450	2	A34169	d3 dopamine recept
42	369.5	17.2	455	2	S28221	alpha-2A-adrenergic
43	369	17.1	458	2	A48392	alpha-2C2 adrener
44	368	17.1	515	2	JC1525	alpha-2C4 adrenoce
45	367.5	17.1	560	2	A38731	alpha-1B-adrenergic
						alpha-1A adrenergic

ALIGNMENTS

RESULT 1

I48231

serotonin receptor 5B - mouse

N:Alternate names: 5-hydroxytryptamine 5B receptor (5HTR-5C)

C:Species: Mus musculus (house mouse)

C>Date: 02-Jul-1996 #sequence_revision 02-Jul-1996 #text_change 05-Nov-1999

C/Accession: I48231

R/Matthes, H.; Boschart, U.; Amlaiky, N.; Grailhe, R.; Plassat, J.L.; Muscatelli, F.; Mat

Mol. Pharmacol. 43, 313-319, 1993

A>Title: Mouse 5-hydroxytryptamine5A and 5-hydroxytryptamine5B receptors define a new fa

A/Reference number: I48231; MUID:93196607; PMID:8450829

A/Accession: I48231

A/Status: preliminary; translated from GB/EMBL/DBD

A/Molecule type: mRNA

A/Residues: 1-370 <RES>

A/Cross-references: EMBL:X69867; NID:g288735; PIDN:CAA49501.1; PID:g288736

C/Superfamily: octopamine receptor type I

C/Keywords: G protein-coupled receptor; glycoprotein; neurotransmitter receptor; transmem

Alignment Scores:

Pred. No.: 4.26e-88 Length: 370
Score: 1452.50 Matches: 298
Percent Similarity: 83.64% Conservative: 19
Best Local Similarity: 78.63% Mismatches: 53
Query Match: 67.43% Indels: 9
DB: 2 Gaps: 3

US-09-976-782-15 (1-1152) x I48231 (1-370)

QY 5 ATGAGAGCGCGTACCTTTTCAGTGGCCACCGCGCGGTTCCTTCCTCGCCCTGGGACCGGAG 64

Db 1 MetGluValSerAsnLeuSerGlyAlaThrProGlyLeuAlaPheProProGlyProGlu 20

QY 65 ACACAGCAGCGGACCCCGGACCCCAACCGGAGAGGATACCTCGGTTCGACCCCGAGCGGC 124

Db 21 SerCysSer-----AspSerProSerGlyArgSerMetGlySerThrProGlyGly 38

QY 125 GCCTCTCTCCGCGGCGGAGGGCGCCCTTCTCTGTCTTCACGGTCTCTGGTGTGTCAGCTG 184

```
Db 39 LeuLeuLeuProGlyArgGluProProPheSerAlaPheThrValLeuValValThrLeu 58
QY 185 CTAGTGTCTGATCGCTGCCACTTCTCTGGAACTCTGCTGGTTCCTGGTTCACCATCCG 244
Db 59 LeuValLeuLeuLeuAlaAlaThrPheLeuTrpAsnLeuLeuValLeuValThrLeu 78
QY 245 CGGTCCTGCTCTCCACCGGTGGCGATTAATCTGGTGGCTTCACCGCGCTCTCGAC 304
Db 79 ArgValArgAlaPheHisArgValProHisAsnLeuValAlaSerThrAlaValSerAsp 98
QY 305 GNACTAGTGGCAGCGCTGCGATGCCACCGACCTGGCGAGTGGAGCTGTGACCGGGGGA 364
Db 99 ValLeuValAlaValLeuValMetProLeuSerLeuValSerGluLeuSerAlaGlyArg 118
QY 365 COTCGCTGCTGGCGGAGCTGTGCCACGTGTGGATCTCTTCGACCGCGGAGCTGT 424
Db 119 ArgTrpGlnLeuGlyArgSerLeuCysHisValTrpIleSerPheAspVal----- 135
QY 425 CTGTGCTGCCCGCGGCTCGGAACTGGGGCCATCGCCCTGGCGCGGACGGGCG 484
Db 136 LeuCysCysThrAlaSerIleTrpAsnValAlaAlaIleAlaLeuAspArgTyrTrpThr 155
QY 485 ATCACACGACCTGCAGCACAGCTGCCACCGCGAGCTGGCAGTGCAGCTGCAGCGCGA 544
Db 156 IleThrArgHisLeuGlnIleValThrLeuValArgSerAlaSerAlaLeuMetIle 175
QY 545 GCGCTCGCCCGGCTGCGCTGCGGCTCATCGCCCTCGCGCTGCTCTTTGCGCGGCG 604
Db 176 AlaIleThrTrpAlaLeuSerAlaLeuIleAlaLeuAlaProLeuLeuPheGlyTrpGly 195
QY 605 GAGGTGTGAGCTCGCTCGCTCAGCGCTGCCAGTGGCGGAGCGGAACTCTATGCGCGC 664
Db 196 GluAlaTyrAspAlaArgLeuGlnArgCysGlnValSerGlnGluProSerTyrAlaVal 215
QY 665 TTCTCCACCGCGCGCTTCCACCTGCGCTTGGCGTGTGCTGCTACCGGGAAG 724
Db 216 PheSerThrCysGlyAlaPheTyrLeuProLeuAlaValValLeuPheValTyrTrpLys 235
QY 725 ATCTACGAGCGCGCCAACTTCTGTTTCGCGCGCGCGCGAGAGCTGTGCTGCGTGGCG 784
Db 236 IleTyrLysAlaAlaLysPheArgPheGlyArgArgAlaValValProLeuPro 255
QY 785 GCACCATGCAAGTGTAGGTCCAGGTAAGGAAAGCACTGATGAGCTGAAGTGGTTC 844
Db 256 AlaThrThrGln-----AlaLysGluAlaProProGluSerGluMetValPhe 271
QY 845 ACGGCACATTGCAAGCAACGCTGCTTCCAGTGCAGCGGAGCTCTGCGCGGAGCAG 904
Db 272 ThrAlaArgArgAlaThrValThrPheGlnThrSerGlyAspSerTrpA-gGluGln 291
QY 905 AAGGAGAGCGCAGCAGCATGATGTGGGAATCTGATGGCGTGTGTTGCTGTGCTGG 964
Db 292 LysGluLysArgAlaAlaMetMetValGlyIleLeuIleGlyValPheValLeuCysTrp 311
QY 965 ATCCCTCTCTCTGACGGAACTCATCAGCCCACTGCTGCTGCTGCGCTGCGCCCATC 1024
Db 312 IleProPhePheLeuThrGluLeuIleSerProLeuCysAlaCysSerLeuProPhe 331
QY 1025 TGGAAAGCATATTCTGTGCTGGCTACTTCCAAATTTCTTTTCAACCCCTGATTTAC 1084
Db 332 TrpLysSerIlePheLeuTrpLeuGlyTyrSerAsnSerPhePheAsnProLeuIleTyr 351
QY 1085 ACAGCTTTTAAAGAACTACAACTGCTTCAAGAGCTCTTTTACTAAGCAGAGA 1141
Db 352 ThrAlaPheAsnLysAsnTyrAsnAsnAlaPheLysSerLeuPheThrLysGlnArg 370
```

RESULT 2

S38744

serotonin receptor 5B - rat

N;Alternate names: 5-hydroxytryptamine receptor 5B (5-HT5B)

C;Species: Rattus norvegicus (Norway rat)

C;Date: 19-May-1994 #sequence_revision 10-Nov-1995 #text_change 05-Nov-1999

C;Accession: S38744; A47472
R.Wieden, W.; Parker, E.M.; Mahle, C.D.; Grisel, D.A.; Nowak, H.P.; Yocca, F.D.; Felder, P.E.S. Lett. 333 25-31, 1993
A;Title: Cloning and characterization of the rat 5-HT(5B) receptor. Evidence that the 5-HT(5B) receptor is a distinct subtype of 5-hydroxytryptamine receptors different from 5-HT(1A) and 5-HT(2A) receptors.
A;Reference number: S38744; MUID:94039744; PMID:8224165
A;Accession: S38744
A;Molecule type: mRNA
A;Residues: 1-369 <WIS>
P.Erlander, M.G.; U.S.A. 90, 3452-3456, 1993
Proc. Natl. Acad. Sci. U.S.A. 90, 3452-3456, 1993
A;Title: Two members of a distinct subfamily of 5-hydroxytryptamine receptors different from 5-HT(1A) and 5-HT(2A) receptors.
A;Reference number: A47472; MUID:93234515; PMID:7682702
A;Accession: A47472
A;Status: preliminary
A;Molecule type: nucleic acid
A;Residues: 1-176, 177-369 <BRL>
A;Cross-references: GB:10073; NID:g310074; PIDN:AAA40616.1; PID:g310075
A;Experimental source: hypothalamus
A;Note: sequence extracted from NCBI backbone (NCBIN:129665, NCBIP:129668)
C;Superfamily: octopamine receptor type I
C;Keywords: G protein-coupled receptor; transmembrane protein

Alignment Scores:

Pred. No.:	1-8e-87	Length:	369
Score:	1443.00	Matches:	236
Percent Similarity:	83.38%	Conservative:	20
Best Local Similarity:	78.10%	Mismatches:	53
Query Match:	66.99%	Indels:	10
DB:	2	Gaps:	4

US-09-976-782-15 (1-1152) x S38744 (1-369)

```
QY 5 ATGAGGCGCTAGCTTTCAGTGGCCACCGCGGCTTGCCTTGCCTTGGACCCGAG 64
Db 1 MetGluValSerAsnLeuSerGlyAlaThrProGlyIleAlaPheProGlyProGlu 20
QY 65 ACCAGCAGCGGACCCCGGACCCCAAGCCGAGAGGATACTCGTTCGACCCCGAGCGG 124
Db 21 SerCysSer-----AspSerProSerSerGlyArgSerMetGlySerThrProGly 38
QY 125 GCGCTCTCCGCGCGGCGGCGGCGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 184
Db 39 LeuIleLeuSerGlyArgGluProPheSerAlaPheThrValLeuValValThrLeu 58
QY 185 CTAGTCTGCTGATCTGCTCCACTTTCCTGTGTGAACCTGCTGTTCGGTTCACATCCG 244
Db 59 LeuValLeuLeuIleAlaAlaThrPheLeuTrpAsnLeuLeuValLeuValThrLeu 78
QY 245 CGGTCCTGCTCTCCACCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCG 304
Db 79 ArgValArgAlaPheHisArgValProHisAsnLeuValAlaSerThrAlaValSerAsp 98
QY 305 GAACTAGTGGCAGCTGGCGATGCCACCGAGCTGGCAGTGCAGCTGCAGCGCGGCGA 364
Db 99 ValLeuValAlaAlaLeuValMetProLeuSerLeuValSerGluLeuSerAlaGlyArg 118
QY 365 CCGTGGCTGTGGCGCGGAGCTGTGCGACGTGTGGATCTCTTCGACCGCGGAGCGCTGT 424
Db 119 ArgTrpGlnLeuGlyArgSerLeuCysHisValTrpIleSerPheAspVal----- 135
QY 425 CTGTGTCGCCCGCGGCTCGGAACTGGGGCCATCGCCCTGGCGCGGCGGCGGCGG 484
Db 136 LeuCysCysThrAlaSerIleTrpAsnValAlaAlaIleAlaLeuAspArgTyrTrpThr 155
QY 485 ATCACACGACCTGCAGCACAGCTGCCACCGCGAGCTGGCAGTGCAGCTGCAGCGCGA 544
Db 156 IleThrArgHisLeuGlnIleValThrLeuValArgSerAlaSerAlaLeuMetIle 175
QY 545 GCGCTCGCCCGGCTGCGCTGCGGCTCATCGCCCTCGCGCTGCTCTTTGCGCGGCG 604
Db 176 AlaThrTrpAlaLeu---SerAlaLeuIleAlaLeuAlaProLeuLeuPheGlyTrpGly 194
QY 605 GAGGTGTGCGAGCTGCGCTCCAGCTCCAGCTGCGGAGTGCAGCGGAACTCTCTATGCGCGC 664
```


A:Title: The mouse 5HT5 receptor reveals a remarkable heterogeneity within the 5HT1D red
A:Reference number: S28058; MUID:93099851; PMID:1464308

A:Accession: S28058

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-357 <PLA>

A:Cross-references: EMBL:Z18278

A:Superfamily: octopamine receptor type I

C:Keywords: G protein-coupled receptor; glycoprotein; neurotransmitter receptor; transmembrane

Alignment Scores:

Pred. No.: 1,18e-59 Length: 357
Score: 1020.50 Matches: 219
Percent Similarity: 70.83% Conservative: 36
Best Local Similarity: 60.83% Mismatches: 94
Query Match: 47.38% Indels: 11
DB: 2 Gaps: 6

US-09-976-782-15 (1-1152) x S28058 (1-357)

QY 65 ACCAGCGGAGCCCGGAGCCCAAGCCGAGGAGGATCTCGTTTCGACCCCGAGCGGC 124

Db 8 ThrSerPheSerLeuSerThrProSerSerLeuGluProAsnArgSerLeuAspThrGlu 27

QY 125 GCGTCTCTCGCGGCGGAGCGCGCCCTTCTCTGCTTTCACGCTCTCGTGTGTGAGCGCTG 184

Db 28 ValLeuArgProSerArg---ProPheLeuSerAlaPheArgValLeuValLeuThrLeu 46

QY 185 CTAGTGTCTGATCGCTGCCACTTTCCTGTGTGGAACCTGCTGGTTCGGTTCACCATCCCG 244

Db 47 LeuGlyPheLeuAlaAlaAlaThrPheThrTrpAsnLeuValLeuAlaThrLeuLeu 66

QY 245 CGGTCCTGCTCTTCCACGCGTGCAGTAACTTGTGGCTCGAGCGCGCTCTCGGAC 304

Db 67 LysValArgThrPheHisArgValProHisAsnLeuValAlaSerMetAlaIleSerAsp 86

QY 305 GAACAGTGTGCGCGCTGGCGATGCCACCGAGCTGGCGAGTGTGCTGCAGCCGCGCGA 364

Db 87 ValLeuValAlaValLeuValMetProLeuSerLeuValHisGluLeuSer---GlyArg 105

QY 365 CTTGGCTGTGGCGGAGCCTGTGCCAGTGTGATCTCTTCGACGCGCGAGCCTGT 424

Db 106 ArgTrpGlnLeuGlyArgLeuGlyGlnLeuTrpIleAlaCysAspVal----- 122

QY 425 CTGTGTGCTGCGCGCGCTGGGAGCTGGCGGCGCATCGCCCTGGCGCGCGAGCGGC 484

Db 123 LeuCysCysThrAlaSerIleTrpAsnValThrAlaIleAlaLeuAspArgTrpSer 142

QY 485 ATCACAGCGCACCTGCAGACAGCTGGCGACCGCGAGCGCGCTCTGTCTCATGATC 544

Db 143 IleThrArgHisLeuGluTrpThrLeuArgThrArgLysArgValSerAsnValMetIle 162

QY 545 GCGCTCGCGCGGTGCGTGGCGCTCATCGCCCTCGCGCGTGTCTTTGCGCGGCGGC 604

Db 163 LeuLeuThrTrpAlaLeuSerThrValIleSerLeuAlaProLeuPheGlyTrpGly 182

QY 605 GAGGTGTGACGCTCGGCTCCAGCGTGCAGTGTGAGCGCGGAACTCTCTATGCGCC 664

Db 183 GluThrTy-SerGluProSerGluGluCysGlnValSerArgGluProSerTrpVal 202

QY 665 TTCTCCACCGCGCGCTTCCACCTGCCCTGGCGTGTGGCGTGTGCTTGTCTACCGGAG 724

Db 203 PheSerThrValGlyAlaPheTyTrpLeuProLeuCysValValLeuPheValTyTrpLys 222

QY 725 ATTACGAGCGGCGAAGTTTCGTTTCGGC---CGCGCGCGGAGAGTGTGTCGCGTGT 781

Db 223 IleTyArgAlaAlaLysPheArgMetGlySerArgLysThrAsnSerValSerProVal 242

QY 782 CCGGCCACCATGCAAGTGTGAGTCCAGGTAAAGGACACCTGATGAGGCTGAGTGGTG 841

Db 243 ProGluAlaValGlu-----ValIysAsnAlaThrGlnHisProGlnMetVal 258

QY 842 TTCACGGCACATTGCAAGCAACCGTGTCTCTCCAGGTGAGCGCGGAGCTCTCGCGGAG 901

Db 259 PheThrValArg---HisAlaThrValThrPheGlnThrGluGlyAspThrTrpArgGlu 277

QY 902 CAGAAGGAGAGCGGAGCCCATGATGTTGGGAATCTGATTTGCGTGTGCTGTCGC 961

Db 278 GlnLysGluGlnArgAlaAlaLeuMetValGlyIleLeuIleGlyValPheValLeuCys 297

QY 962 TGGATCCCTCTTCTCTGACGGAACCTCATCAGCCCACTGTGCTGCGACCTGCGCCCC 1021

Db 298 TrpPheProPhePheValThrGluLeuIleSerProLeuCysSerTrpAspValProAla 317

QY 1022 ATTCGAAAAGCATATTTCTGTGGCTGTGCTACTCTCAATCTTCTTCAACCCCTGAT 1081

Db 318 IleTrpLysSerIlePheLeuTrpLeuGlyTyTrpAsnSerPhePheAsnProLeuIle 337

QY 1082 TACACAGCTTTTAAACAGAACTACAACTGCTTCAAGAGCTCTTTTAAAGCAGAGA 1141

Db 338 TyrThrAlaPheAsnArgSerTySerAlaPheLysValPhePheSerLysGlnGln 357

RESULT 5

B47472

5-hydroxytryptamine 5 alpha receptor - rat

C:Species: Rattus norvegicus (Norway rat)

C:Date: 21-Jan-1994 #sequence_revision 18-Nov-1994 #text_change 05-Nov-1999

C:Accession: B47472

R:Erlander, M.G.; Lovenberg, T.W.; Baron, B.M.; de Lecea, L.; Danielson, P.E.; Racke, M.

Proc. Natl. Acad. Sci. U.S.A. 90, 3452-3456, 1993

A:Title: Two members of a distinct subfamily of 5-hydroxytryptamine receptors differ in

A:Reference number: A47472; MUID:93234515; PMID:7682702

A:Accession: B47472

A>Status: preliminary

A:Molecule type: nucleic acid

A:Residues: 1-357 <EURL>

A:Cross-references: GB:110072; NID:G310072; PIDN:AAA40615.1; PID:G310073

A:Experimental source: hypothalamus

A:Note: sequence extracted from NCBI backbone (NCBIN:129674, NCBI:129677)

C:Superfamily: octopamine receptor type I

C:Keywords: G protein-coupled receptor; transmembrane protein

Alignment Scores:

Pred. No.: 2,94e-59 Length: 357
Score: 1014.50 Matches: 216
Percent Similarity: 71.27% Conservative: 37
Best Local Similarity: 60.85% Mismatches: 81
Query Match: 47.10% Indels: 21
DB: 2 Gaps: 7

US-09-976-782-15 (1-1152) x B47472 (1-357)

QY 110 TCGACCCCGAGCGCGCTCTCTCGCGGCGGAGCGCGCC----- 151

Db 13 SerThrProSer---ThrLeuGluProAsnArgSerLeuAspThrGluAlaLeuArgThr 31

QY 152 -----TTCTCTGTCTTCAGCGTCTCGTGTGTGACGCTGTAGTGTGCTGATC 199

Db 32 SerGlnSerPheLeuSerAlaPheArgValLeuValLeuThrLeuLeuGlyPheLeuAla 51

QY 200 GCTGCCACTTCTCTGTGGAACTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 259

Db 52 AlaAlaThrPheThrTrpAsnLeuLeuValLeuAlaThrIleLeuArgValArgThrPhe 71

QY 260 CACCGGTGCGCGCATAACTTGTGGCTCGACCGCGCTCTCGACGAACTAGTGGGAGCG 319

Db 72 HisArgValProHisAsnLeuValAlaSerMetAlaIleSerAspValLeuValAlaVal 91

QY 320 CTGGCGATGCCCGGAGCTGCGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 379

Db 92 LeuValMetProLeuSerLeuValHisGluLeuSer---GlyArgArgTrpGlnLeuGly 110

QY 380 CGGAGCTGTGCGACGTGTGATCTCTCTCGACGCGGAGCTGTGTGTGTGTGTGTGTGTGT 439

Db 111 ArgArgLeuCysGlnLeuTrpIleAlaCysAspVal-----LeuCysCysThrAla 127


```

QY 740 AAGTTTCGTTC-----GCCGCCGC---CGGAGAGCTGTG 772
Db 236 ArgSerArgIleLeuLysGlnThrProAsnArgThrGlyLysArgLeuThrArgAlaGln 255
QY 773 CTG-----CCGTGCGGCCACCATGTCAGGTGAGGTCCAAAGTAAG 814
Db 256 LeuIleThrAspSerProGlySerThrSerValThrSerIleAsnSerArgValPro 275
QY 815 GAAGCACCCTGATAGGCT---GAAGTGGTGTTCACCGGCACATTCGAAAGCAACGTGTCC 871
Db 276 AspValProSerGluSerGlySerProValTyValAsnGlnValLysValArgValSer 295
QY 872 TTCAGAGTACGGGGACTCTCGCGCGGAGCAGAAG-----GAGAGG 913
Db 296 -----AspAlaLeuLeuGluLysLysLysLeuMetAlaAlaArgGluArg 310
QY 914 CGAGCAGCCATGATGGTGGGAATCTGATGGCGTGTGTGCTGTGTCGATCCCTTC 973
Db 311 LysAlaThrLysThrLeuGlyIleLeuGlyAlaPheIleValCysTrpLeuProPhe 330
QY 974 TTCCTGACGGAACTCATCAGCCCATCTGT-----GCTGCGAGCCTGCCGCCCATCTGG 1027
Db 331 PheIleIleSerLeuValMetProIleCysLysAspAlaCysTrpPheHisLeuAlaIle 350
QY 1028 AAAAGCATATTTCTGCTGGCTGCTACTCCATCTCTTTCAACCCCTGATTTACACA 1087
Db 351 PheAspPhePheThrTrpLeuGlyTyTrpLeuAsnSerLeuIleAsnProIleIleTyThr 370
QY 1088 GCTTTTACAGAACTACAAACATGCTTCAGAGCCTC 1126
Db 371 MetSerAsnGluAspPheGlnAlaPheHisLysLeu 383

RESULT 8
S58126
serotonin receptor 1-like - rabbit
N;Alternate names: 5-hydroxytryptamine receptor 1-like (5-HTR)
C;Species: Oryctolagus cuniculus (domestic rabbit)
C;Date: 13-Jan-1996 #sequence_revision 01-Mar-1996 #text_change 05-Nov-1999
C;Accession: S58126
R;Wurch, T.; Cathala, C.; Palmer, C.; Valentin, J.P.; John, G.; Colpaert, F.C.; Pauwels,
submitted to the EMBL Data Library, July 1995
A;Reference number: S58126
A;Accession: S58126
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-390 <WUR>
A;Cross-references: EMBL:X89731; NID:g1052771; PIDN:CAA61893.1; PID:g1052772
C;Superfamily: vertebrate rhodopsin
C;Keywords: neurotransmitter receptor

Alignment Scores:
Pred. No.: 1.08e-21 Length: 390
Score: 444.00 Matches: 117
Percent Similarity: 51.30% Conservative: 60
Best Local Similarity: 33.91% Mismatches: 132
Query Match: 20.61% Indels: 36
DB: 2 Gaps: 11

US-09-976-782-15 (1-1152) x S58126 (1-390)

QY 167 GTCTGTGTGAGCGTGTAGTCTGCTGATCGCTGCCACTTTCCTGTGGAACCTGCTG 226
Db 50 ValLeuLeuValLeuLeuAlaLeuPheThrLeuAlaThrThrLeuSerAsnAlaPhe 69
QY 227 GTTCCGGTACCATCCCGGGTCCGCTTCACCGCGTGCAGTACTTGTGGTGGCC 286
Db 70 ValValAlaThrValTyArgThrArgLysLeuHisThrProAlaAsnTyTrpLeuAla 89
QY 287 TCGACGCGCTCTCGGACGAACTAGTGGCGCTGCGGATGCCACCGAGCTGGCGAGT 346
Db 90 SerLeuAlaValThrAspLeuValSerIleLeuValMetPro-----IleSerThr 107
QY 347 GAGCTGTCGACGGGAGCTGGCTGCTGGCGCGGAGCCTGTGCCACGTGTGATCTCC 406

```

```

Db 108 MetTyThrValThrGlyArgTrpThrLeuGlyGlnValValCysAspLeuTrpLeuSer 127
QY 407 TTCGACGCGGAGCCTGTCTGTGTCGCCCGCCGCTCGGAAAGCTGGCGCCATCGCC 466
Db 128 SerAsp-----IleThrCysCysThrAlaSerIleMetHisLeuCysValIleAla 144
QY 467 CTGGGCGCGGACCGGCGCATCACACGGCACCTGCGACGACCGCTCGGCACCGCCGAGCCGC 526
Db 145 LeuAspArgTyTrpAlaIleThrAspAlaValGluTyTrpSerAlaLysArgThrProLys 164
QY 527 GCCTGTGCTCATGATCGCTCGCCCGGGTGGCTCGCGCTCATGCCCTCGCCGCGC 586
Db 165 ArgAlaAlaIleMetIleArgLeuValTrpValPheSerIleCysIleSerLeuPro 184
QY 587 CTGCTCTTTGGCGGCGGAGGTGTGCGACGCTCGGCTCCAGCGCTGCGAGGTGAGCGG 646
Db 185 ---PhePheTrpArgGlnAlaLysAlaGluGluValSerGluCysLeuValAsnThr 203
QY 647 GAA---CCCTCCTATGCCGCTTCTCCACCGCGCGGCTTCCACCTCGCGCTTGGCGTG 703
Db 204 AspHisValLeuTyThrValTySerThrValGlyAlaPheTyLeuProThrLeuLeu 223
QY 704 GTGCGCTTTGTCTACCGGAGATCTACGAGCGCGCAAGTTTCGTTC----- 751
Db 224 LeuIleAlaLeuTyGlyArgIleTyTrpValGluAlaArgSerArgIleLeuLysGlnThr 243
QY 752 -----GCCGCCGC---CGGAGAGCTGTGCTG-----CCG 778
Db 244 ProAsnArgThrGlyLysArgLeuThrArgAlaGlnLeuIleThrAspSerProGlySer 263
QY 779 TTGCGCGCCACCATGACAGTCCAGGTCCAAAGTAAGGACACCTGATGAGCT---GAA 835
Db 264 ThrThrSerValThrSerIleAsnSerArgAlaProAspValProSerGluSerGlySer 283
QY 836 GTGCTGTTCACGGCACATTGCAAGCAACCGTGTCTTCCAGGTGAGCGGGACTCTCTGG 895
Db 284 ProValTyValAsnGlnValLysValArgValSer-----AspAlaLeu 298
QY 896 CGGAGCAGAGAG-----GAGAGCGGAGCAGCCATGATGTGGGAATT 937
Db 299 LeuGluLysLysLysLeuMetAlaAlaArgGluArgLysAlaThrLysThrLeuGlyIle 318
QY 938 CTGATTGCGCTGTTGTGCTGTGCTGATCCCTCTTCTTCGCGGAACTCATCAGCCCA 997
Db 319 IleLeuGlyValPheIleValCysTrpLeuProPhePheIleIleSerLeuValMetPro 338
QY 998 CTCTGT-----GCCTGCGAGCCTGCCGCCCATCTCGAAAGACATATTCTGTGCTGGC 1051
Db 339 IleCysLysAspAlaCysTrpPheHisGlnAlaIlePheAspPhePheThrTrpLeuGly 358
QY 1052 TACTCCCAATTTCTTCTCAACCCCTGATTTACAGCTTTTAAACAAGAACTACAACAAT 1111
Db 359 TyrValAsnSerLeuIleAsnProIleIleTyTrpMetSerAsnGluAspPheLysGln 378
QY 1112 GCCTTCAAGAGCCTC 1126
Db 379 AlaPheHisLysLeu 383

RESULT 9
B30341
G protein-coupled receptor RDC4 - dog
C;Species: Canis lupus familiaris (dog)
C;Date: 07-Jun-1990 #sequence_revision 07-Jun-1990 #text_change 05-Nov-1999
C;Accession: B30341; S12822
R;Libert, F.; Parmentier, M.; Lefort, A.; Dinsart, C.; Van Sande, J.; Maenhaut, C.; Simor
Science 244, 569-572, 1989
A;Title: Selective amplification and cloning of four new members of the G protein-coupled
A;Reference number: A30341; MUID:89242119; PMID:2541503
A;Accession: B30341
A;Status: nucleic acid
A;Molecule type: mRNA
A;Residues: 1-377 <LIB>

```


A;Cross-references: EMBL:X14049; NID:g900; PIDN:CAA32207.1; PID:g901
 R;Libert, F.; Parmentier, M.; Lefort, A.; Dumont, J.B.; Vassart, G.
 Nucleic Acids Res. 18, 1916, 1990
 A;Title: Complete nucleotide sequence of a putative G protein coupled receptor: RDC4.
 A;Reference number: S12822; MUID:90245610; PMID:2159630
 A;Accession: S12822
 A;Status: translation not shown
 A;Molecule type: mRNA
 A;Residues: 1-377 <L12>
 A;Cross-references: EMBL:X14049; NID:g900; PIDN:CAA32207.1; PID:g901
 C;Genetics:
 A;Gene: RDC4
 A;Superfamily: octopamine receptor type I
 C;Keywords: G protein-coupled receptor; glycoprotein; transmembrane protein

Alignment Scores:
 Pred. No.: 1,7e-21 Length: 377
 Score: 441.00 Matches: 121
 Percent Similarity: 47.10% Conservative: 66
 Best Local Similarity: 30.48% Mismatches: 142
 Query Match: 20.47% Indels: 68
 DB: 2 Gaps: 11

US-09-976-782-15 (1-1152) x B30341 (1-377)

QY 8 GAGCGCGTACGCTTTCAGTGGCCACCGCGGGTGGCCCTGGCCGCGGAGCC 67
 Db 14 GluAlaSerAsnArgSerLeuAsnAlaThrGluThrProGluAlaTrpGlyProGluThr 33
 QY 68 AGCAGCGGACCGCGGACCCCAAGCCCGAGAGGATACCTCGTTGACCCCGCGCGCC 127
 Db 33 ----- 33

QY 128 GTCTGCGGCGGCGAGGCGCGCCCTCTCTGCTTCACGGTCTGCTGTGCTGCTGCTA 187
 Db 34 -----LeuGlnAlaLeuLysIleSerLeuAlaLeuLeu 45
 QY 188 GTGCTGCTGATCGCTGCCACTTTCGTTGGACCTGCTGCTCCGGTCCACCTCCCGCG 247
 Db 46 SerIleIleThrMetAlaThrAlaLeuSerAsnAlaPheValLeuThrThrIlePheLeu 65
 QY 248 GTCCGTCGCTTCACCGCGTCCGCGATACCTGTGTCGCTCGACGCGCTCTCGGACGAA 307
 Db 66 ThrArgLysLeuHisThrProAlaAsnTyrLeuIleGlySerLeuAlaMetThrAspLeu 85
 QY 308 CTAGTGGCAGCGTGGCGATGCCACCGAGCTCGGAGTGTGCTGCGACCGCGGACGT 367
 Db 86 LeuValSerIleLeuValMetProIleSerIleAla-----TyrThrThrArgThr 103
 QY 368 CGGCTGCTGGCGGCGAGCGCTGCGCGTCCGCTGCTGCTGCTGCTGCTGCTGCTGCTG 427
 Db 104 TrpSerPheGlyGlnIleLeuLysAspIleTrpLeuSerSerAsp-----IleThr 120
 QY 428 TGCTGCGCGCGCTCGGAAAGTGGCGGCGCATCGCCCTGGCGCGCGCGGCGCATC 487
 Db 121 CysCysThrAlaSerIleLeuHisLeuValIleAlaLeuAspArgTyrTrpAlaIle 140
 QY 488 ACAGCGACCTGCGACACGCTGCGGACCCCGCGCGCGCGCTGCTGCTGCTGCTGCTG 547
 Db 141 ThrAspAlaLeuGluTyrSerLysArgThrAlaGlyArgAlaAlaValMetIleAla 160
 QY 548 CTGCGCGGCGTGGCTGCGGCGCTCATCGCCCTCGCGCGCTGCTCTTTGGCGGCGG 607
 Db 161 ThrValTrpValIleSerIleCysIleSerIleProPro---LeuPheTrpArgGlnAla 179
 QY 608 GTGTGCGACCTCGGCTCCAGCGCTGCCAGTG---AGCGGGAACCTCTCTATGCGGCC 664
 Db 180 LysAlaGlnGluAspMetSerAspCysGlnValAsnThrSerGlnIleSerTyrThrIle 199
 QY 665 TTCTCCACCGCGCGCTCCACCTCGCGCTGGCGTGGCTGCTGCTGCTGCTGCTGCTG 724
 Db 200 TyrSerThrCysGlyAlaPheTyrIleProSerValLeuLeuIleLeuTyrGlyArg 219

QY 725 ATCTACGAGCGCGCAAGCTTTGCT-----TTCCGCGCGCGCGCG 763
 Db 220 lIeTyrValAlaAlaArgAsnArgIleLeuAsnProProSerLeuTyrGlyLysArgPhe 239
 QY 764 AGAGCTGTGTCGCGTGG-----CCGGCC 787
 Db 240 ThrThrAlaGlnLeuIleThrGlySerAlaGlySerLeuCysSerLeuSerProSer 259
 QY 788 ACCATGTCAGGTGAGGTCCAAGGTAAGCAAGCACCTGATGAGGCTCAAGTGTGTTACG 847
 Db 260 LeuGlnGluGluArgSerHisAlaAlaGlyProPro-----LeuPhePhe 274
 QY 848 GCACATTGCAAGCAACGGTGTCTTCCAGGTGAGCGGGAGCTCTCGCGGAGAGCAG---- 904
 Db 275 AsnHisValGlnValLysLeuAla-----GluGlyValLeuGluArgLysArgIle 291
 QY 905 -----AGGAGAGCGGAGCAGCCATGATGGTGGGAATCTGATGGCGTGTG 955
 Db 292 SerAlaAlaArgGluArgLysAlaThrLysThrLeuGlyIleLeuGlyAlaPheIle 311
 QY 956 CTGTGCTGATCCCTCTTCTCTGACGGAACCTCATCAGCCCACTCTGT-----GCTGC 1009
 Db 312 ValCysTrpLeuProPhePheValAlaSerLeuValLeuProIleCysArgAlaSerCys 331
 QY 1010 AGCTGCTCCCGCCATCTGGAAGCATATTTCTGTGCTTGGCTTACTCCAAATCTTCTTC 1069
 Db 332 TrpLeuHisProAlaLeuPheAspPheThrTrpLeuGlyTyrLeuAsnSerLeuIle 351
 QY 1070 ACCCGCTGATTTACACAGCTTTTACAGAACTCAACAATGCCCTTCAAG 1120
 Db 352 AsnProIleIleTyrThrValPheAsnGluGluPheArgGlnAlaPheGln 368

RESULT 10

S54153

serotonin receptor 1B - Chinese hamster

N;Alternate names: 5-hydroxytryptamine receptor 1B (5-HTR1B)

C;Species: Cricetulus griseus (Chinese hamster)

C;Date: 08-Jul-1995 #sequence_revision 21-Jul-1995 #text_change 05-Nov-1999

C;Accession: S54153

R;Wurch, T.; Palmier, C.; Colpaert, F.C.; Pauwels, P.J.

submitted to the EMBL Data Library, April 1995

A;Description: Molecular cloning and expression of a Chinese hamster lung fibroblast

A;Reference number: S54153

A;Accession: S54153

A;Status: preliminary

A;Molecule type: mRNA

A;Residues: 1-386 <WUR>

A;Cross-references: EMBL:X86458; NID:g790985; PIDN:CAA60175.1; PID:g790986

C;Superfamily: vertebrate rhodopsin

C;Keywords: neurotransmitter receptor

Alignment Scores:

Pred. No.: 1,97e-21 Length: 386
 Score: 440.00 Matches: 120
 Percent Similarity: 51.01% Conservative: 56
 Best Local Similarity: 34.78% Mismatches: 133
 Query Match: 20.43% Indels: 36
 DB: 2 Gaps: 12

US-09-976-782-15 (1-1152) x S54153 (1-386)

QY 167 GTCTGCTGCTGAGCTGCTGTAGTGTGCTGCTGATCGCTGCCACTTCTCTGTGGAACTGCTG 226

Db 46 ValLeuValAlaLeuLeuAlaLeuLeuAlaLeuLeuAlaThrThrLeuSerAsnAlaPhe 65

QY 227 GTTCCGCTCACCATCCCGGGTCCGTGCTTCCACCGCGTCCCGGCAACTTGTGGGCC 286

Db 66 ValIleAlaThrValTyrArgThrArgLysLeuHisThrProAlaSerTyrLeuIleAla 85

QY 287 TCGACGCGCTCTCGGACGAACTAGTGGCAGCGCTGCGCATGCCAGCGCTGCGGAGT 346

Db 86 SerLeuAlaValThrAspLeuValSerIleLeuValMetProValSerThrMetTyr 105

US-09-976-782-15 (1-1152) x T42203 (1-445)

Qy	47	CTTGCCCTGGACCCGAGACGACGAGCGGACCCGGGACCCCAAGCCGCGAGAGGATACCT	106
Db	7	LeuAsnLeuThrAlaSerProSerGlnProAlaThr	21
Qy	107	GGTTGCACCCGAGCGGCGCGTCTCTG	154
Db	22	SerAlaValAlaArgGlyThrHisLeuValAspGlnPheProAlaHisalaGluIlePhe	41
Qy	155	TCT	196
Db	42	SerAspIleGluArgProLeuValGlnThrValIleLeuAlaSerValLeuValLeu	61
Qy	197	ATCGCTGCACCTTCTCTGTGGAACTGCTGTTCGGTCCACCATCCCGGGTCCGTGCC	256
Db	62	IleLeuSerCysPheIleGlyAsnLeuPheValIleLeuAlaIleMetGluArgAsp	81
Qy	257	TTCCAC--CGCGTCCGCATAACTGTGTGGCTCGACGGCGCGTCTCGGACGAACCTAGTG	313
Db	82	LeuArgGlyArgProGlnTyrIleLeuPheSerLeuAlaValAlaAspLeuVal	101
Qy	314	GCAGCGCTGGCATGCCACGAGCTGGCGAGTGCAGCTGCACCGGGCGAGTCCGGCTG	373
Db	102	GlyMetIleValThrPro--LeuGlyAlaThrPheThrValThrGlyThrTrpAsn	119
Qy	374	CTGGGCGGAGCTGTGCCACCTGTGGATCTCTTCGACGCGGAGCCCTGTCTGTGCTGC	433
Db	120	LeuGlyValValValCysAspPheTrpIleSerValAspVal--LeuValCys	136
Qy	434	CCGCGCGCTCGGAAAGTGGCGGCATCGCGCCCTGGCGCGGACGGGCGCATCACACGG	493
Db	137	ThrAlaSerIleLeuHisLeuValAlaIleAlaLeuAspArgTyrTrpSerIleThrAsp	156
Qy	494	CACCTGCAGCACACGCTGGCGACCCGCGCGCTCGTGTCTATGATCGCGTCCGCG	553
Db	157	IleCysTyrValGlnAsnArgThrProLysArgIleThr--LeuMetLeuAlaValIle	175
Qy	554	CGGTGCGCTGCGCGCTCATCCGCTCGCGCGCTGCTCTTTGGCGGGCGAGGTGTC	613
Db	176	TrpPheThrSerLeuLeuIleSerLeuAlaProPheAlaGlyTrpLysAspGluGlyPhe	195
Qy	614	GACGCTCGCTCCAGCGC--TCCAGGTGAGCGCGGAAACCTCTATGCCGCG	664
Db	196	SerAspArgValLeuLysSerHisValCysLeuIleSerGlnGlnIleSerTyrGlnVal	215
Qy	665	TTCTCCACCGCGCGCCTTCACCTGCGCGTGGCGTGGCGCGTTCCTACCGGAG	724
Db	216	PheSerThrAlaThrAlaPheTyrIleProLeuIleAlaIleIleCysValTyrTrpLys	235
Qy	725	ATCTACGAGGCGGCAAGTTTCGTTTCGCGCGC--CGCCGGAGAGCTGTGCTG	775
Db	236	IleMetArgAlaAlaLysArgPheLysArgGluArgAspArgThrValIleArg	255
Qy	776	CGTGTGCCG--GCCACCATCAGGTGAGGTCCAAG--	808
Db	256	ProProAspAlaIleAspGluLysAlaMetMetProLysSerLysCys	275
Qy	808	--	808
Db	276	ProLeuProProAlaValValIleSerAspIleGlnAlaAsnGlyGlyThrGlyGlyLys	295
Qy	809	--GTAAGGAGACACCT--GAT	826
Db	296	ThrAsnSerIleLysAsnProProArgHisAsnGluSerSerSerAlaSerGluGlu	315
Qy	827	GAGGCTGAAGTGGTTCACGGCACATTGCCAAGCAACGCTGCTCCTCCAGGTG--	880
Db	316	GluArgThrMetThrGlnThrAsnHisAlaProGlyAspValThrThrGlnGluThrLys	335
Qy	881	--AGCGGG	886

RESULT 15

A53279
serotonin receptor 1D - human
N:Alternate names: 5-hydroxytryptamine receptor 1D (5-HT1D); 5HT-1D alpha receptor
C:Species: Homo sapiens (man)
C>Date: 12-May-1994 #sequence_revision 12-May-1994 #text_change 05-Nov-1999
C:Accession: A53279; A44136
R:Hamblin, M.W.; Metcalf, M.A.
Mol. Pharmacol. 40, 143-148, 1991
A:Title: Primary structure and functional characterization of a human 5-HT-1D-type sero
A:Reference number: A53279; MUID:91342595; PMID:1652050
A:Accession: A53279
A:Molecule type: DNA
A:Residues: 1-377 <HAM>
A:Cross-references: GB:M89955; NID:g177771; PIDN:AAA35491.1; PID:g177772
R:Weinstein, R.L.; Zgonick, J.M.; Macchi, M.J.; Branche, T.A.; Hartig, P.R.
Proc. Natl. Acad. Sci. U.S.A. 89, 3630-3634, 1992
A:Title: Human serotonin 1D receptor is encoded by a subfamily of two distinct genes: 5-
A:Reference number: A44136; MUID:92228840; PMID:1565658
A:Accession: A44136
A:Molecule type: mRNA
A:Residues: 1-377 <WEI>
A:Cross-references: GB:M81589; NID:g338023; PIDN:AAA60315.1; PID:g338024
C:Genetics:
A:Gene: GDB:HTR1D
A:Cross-references: GDB:132416; OMIM:182133
A:Map position: lp36.3-lp34.3
A:Introns: #status absent
C:Superfamily: octopamine receptor type I
C:Keywords: G protein-coupled receptor; glycoprotein; neurotransmitter receptor; transmem

Alignment Scores:	
Pred. No.:	Length:
Score: 430.00	Matches: 377
Percent Similarity: 48.96%	Conservative: 70
Best Local Similarity: 30.73%	Mismatches: 156
Query Match: 19.96%	Indels: 40
DB:	Gaps: 12

US-09-976-782-15 (1-1152) x A53279 (1-377)

Qy	53	CTGGGACCGAGACGACGACGCGACCGACCCCA-----AGCCGAGAGGATA	103
		1 ::::	118
		MetSerProLeuAsnGlnSerAlaGluGlyLeuProGlnGluAlaSerAsnArgSerLeu	20
		1 ::::	156
		104 CTCGGTTCGACCCGACGCGCGCTCTCCGCGCGCGAGCGCGCTCTCTGTCCTC	163
		1 ::::	12
		21 AsnAlaThrGluThrSerGluAlaTrpAspProArgThrLeuGlnAlaLeuLysIleSer	40
		164 ACGTCTCGGTGTGAGCTGCTAGTCTGATCGCTCCACTTTCCTGTGGAACCTG	223
		1 ::::	118

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - protein search, using frame_plus_n2p model

Run on: August 31, 2004, 20:21:44 ; Search time 160 Seconds

(without alignments)
4530.426 Million cell updates/sec

Title: US-09-976-782-15

Perfect score: 2154

Sequence: 1 cgcacgagccgctagcc.....aagcagagatgaacacaggg 1152

Scoring table:

BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 1297172 seqs, 314612898 residues

Total number of hits satisfying chosen parameters: 2594344

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+n2p.model -DEV=xlp
-Q=/cgn2_1/USPRO_spool_p/US09976782/runat_31082004_115953_10443/app_query.fasta_1.1351
-DB=Published Applications AA -QFMT=fastan -SUFFIX=n2p.rapb -MINMATCH=0.1
-LOOPCL=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62
-TRANS=human40.cdi -LIST=45 -DOCLIGN=200 -THR_SCORE=pct -THR_MAX=100
-THR_MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0
-MAXLEN=2000000000 -USER=US09976782@cgn_1_1_225_@runat_31082004_115953_10443
-NCPU=6 -ICPU=3 -NO_MMAP -LARGESQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100
-LONGLOG -DEV_TIMEOUT=120 -WARN_TIMEOUT=130 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5
-FGAPOP=6 -FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Published Applications AA:

1: /cgn2_6/ptodata/2/pubppa/US07_PUBCOMB.pcp.*
2: /cgn2_6/ptodata/2/pubppa/PCT_NEW_PUB.pcp.*
3: /cgn2_6/ptodata/2/pubppa/US06_NEW_PUB.pcp.*
4: /cgn2_6/ptodata/2/pubppa/US06_PUBCOMB.pcp.*
5: /cgn2_6/ptodata/2/pubppa/US07_NEW_PUB.pcp.*
6: /cgn2_6/ptodata/2/pubppa/PCTUS_PUBCOMB.pcp.*
7: /cgn2_6/ptodata/2/pubppa/US08_NEW_PUB.pcp.*
8: /cgn2_6/ptodata/2/pubppa/US08_PUBCOMB.pcp.*
9: /cgn2_6/ptodata/2/pubppa/US09A_PUBCOMB.pcp.*
10: /cgn2_6/ptodata/2/pubppa/US09B_PUBCOMB.pcp.*
11: /cgn2_6/ptodata/2/pubppa/US09C_PUBCOMB.pcp.*
12: /cgn2_6/ptodata/2/pubppa/US09_NEW_PUB.pcp.*
13: /cgn2_6/ptodata/2/pubppa/US10A_PUBCOMB.pcp.*
14: /cgn2_6/ptodata/2/pubppa/US10B_PUBCOMB.pcp.*
15: /cgn2_6/ptodata/2/pubppa/US10C_PUBCOMB.pcp.*
16: /cgn2_6/ptodata/2/pubppa/US10_NEW_PUB.pcp.*
17: /cgn2_6/ptodata/2/pubppa/US60_NEW_PUB.pcp.*
18: /cgn2_6/ptodata/2/pubppa/US60_PUBCOMB.pcp.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	ID	Description

1	1946	90.3	379	10	US-09-976-782-16	Sequence 16, Appl
2	1861.5	86.4	370	16	US-10-333-946-15	Sequence 15, Appl
3	1860.5	86.4	370	10	US-09-823-187-10	Sequence 10, Appl
4	1857.5	86.2	370	10	US-09-823-187-12	Sequence 12, Appl
5	1856.5	86.2	372	12	US-10-311-671-8	Sequence 8, Appl
6	1819.5	84.5	380	10	US-09-954-342-30	Sequence 30, Appl
7	1819.5	84.5	380	10	US-09-823-187-56	Sequence 56, Appl
8	1453.5	67.5	370	10	US-09-823-187-57	Sequence 57, Appl
9	1453.5	67.5	370	10	US-09-954-342-56	Sequence 56, Appl
10	1453.5	67.5	370	10	US-09-976-782-101	Sequence 101, Appl
11	1452.5	67.4	370	10	US-09-823-187-58	Sequence 58, Appl
12	1452.5	67.4	370	10	US-09-954-342-57	Sequence 57, Appl
13	1452.5	67.4	370	10	US-09-976-782-102	Sequence 102, Appl
14	1452.5	67.4	370	14	US-10-109-532A-2	Sequence 2, Appl
15	1443	67.0	369	10	US-09-823-187-59	Sequence 59, Appl
16	1443	67.0	369	10	US-09-976-782-103	Sequence 103, Appl
17	1296.5	60.2	606	14	US-10-017-161-698	Sequence 698, Appl
18	1296.5	60.2	606	15	US-10-292-798-610	Sequence 610, Appl
19	1063.5	49.4	357	10	US-09-823-187-60	Sequence 60, Appl
20	1063.5	49.4	357	10	US-09-954-342-58	Sequence 58, Appl
21	1063.5	49.4	357	10	US-09-976-782-104	Sequence 104, Appl
22	1063.5	49.4	357	14	US-10-225-567A-446	Sequence 446, Appl
23	1063.5	49.4	357	14	US-10-345-680-38	Sequence 38, Appl
24	1063.5	49.4	357	15	US-10-352-684A-16	Sequence 16, Appl
25	1058.5	49.1	357	11	US-09-826-509-447	Sequence 447, Appl
26	1022.5	47.5	357	10	US-09-954-342-60	Sequence 60, Appl
27	1014.5	47.1	357	10	US-09-954-342-59	Sequence 59, Appl
28	1014.5	47.1	357	10	US-09-976-782-105	Sequence 105, Appl
29	647.5	30.1	212	9	US-09-750-373-18	Sequence 18, Appl
30	501	23.3	99	9	US-09-750-373-17	Sequence 17, Appl
31	448	20.8	390	13	US-10-005-010-4	Sequence 4, Appl
32	448	20.8	390	14	US-10-225-567A-4	Sequence 4, Appl
33	448	20.8	390	14	US-10-118-804-15	Sequence 15, Appl
34	448	20.8	390	14	US-10-318-661-20	Sequence 20, Appl
35	448	20.8	390	15	US-10-366-288-40	Sequence 40, Appl
36	447	20.8	390	11	US-09-826-509-427	Sequence 427, Appl
37	430.5	20.0	390	14	US-10-166-101-6	Sequence 6, Appl
38	430	20.0	377	13	US-10-005-010-2	Sequence 2, Appl
39	430	20.0	377	14	US-10-225-567A-6	Sequence 6, Appl
40	430	20.0	377	14	US-10-118-804-14	Sequence 14, Appl
41	430	20.0	377	14	US-10-318-661-19	Sequence 19, Appl
42	430	20.0	422	14	US-10-166-101-3	Sequence 3, Appl
43	430	20.0	422	14	US-10-118-804-13	Sequence 13, Appl
44	429	19.9	377	11	US-09-826-509-429	Sequence 429, Appl
45	428	19.9	422	14	US-10-318-661-10	Sequence 10, Appl

ALIGNMENTS

RESULT 1
US-09-976-782-16
; Sequence 16, Application US/09976782
; Publication No. US20030190715A1
; GENERAL INFORMATION:
; APPLICANT: Grosse et al
; TITLE OF INVENTION: No. US20030190715A1e1 Proteins and Nucleic Acids Encoding Same
; FILE REFERENCE: 21402-157
; CURRENT APPLICATION NUMBER: US/09/976,782
; PRIOR FILING DATE: 2001-10-12
; PRIOR APPLICATION NUMBER: 60/240,113
; PRIOR FILING DATE: 2000-10-12
; PRIOR APPLICATION NUMBER: 60/240,662
; PRIOR FILING DATE: 2000-10-16
; PRIOR APPLICATION NUMBER: 60/240,732
; PRIOR FILING DATE: 2000-10-16
; PRIOR APPLICATION NUMBER: 60/240,625
; PRIOR FILING DATE: 2000-10-16
; PRIOR APPLICATION NUMBER: 60/240,703
; PRIOR FILING DATE: 2000-10-16
; PRIOR APPLICATION NUMBER: 60/241,190
; PRIOR FILING DATE: 2000-10-16
; PRIOR APPLICATION NUMBER: 60/240,637
; PRIOR FILING DATE: 2000-10-16

Alignment Scores:			
Pred. No.:	1.96e-107	Length:	370
Score:	1861.50	Matches:	370
Percent Similarity:	97.63%	Conservative:	0
Best Local Similarity:	97.63%	Mismatches:	0
Query Match:	86.42%	Indels:	9
DB:	16	Gaps:	3
US-09-976-782-15 (1-1152) x US-10-333-946-15 (1-370)			
QY	5	ATGAGCGCGCTAGCCTTTCAGTGGCCACCGCGCGCTTGCCTTGCCTGGACCCGAG	64
Db	1	MetGluAlaAlaSerLeuSerValAlaThrAlaGlyValAlaLeuAlaLeuGlyProGlu	20
QY	65	ACCAGACGGGACCCGGGACCCAAAGCCCGAGAGGATCTCGGTTCGACCCCGAGCGGC	124
Db	21	ThrSerSer-----GlyThrProSerProArgGlyLeuGlySerThrProSerGly	38
QY	125	GGCGTCTCGCGGCGGAGGGCGGCCCTTCTCTGTTCACGGTCTCTGGTGGTACGCTG	184
Db	39	AlaValLeuProGlyArgGlyProProPheSerValPheThrValLeuValValThrLeu	58
QY	185	CTAGTGCTGCTCATCGCTGCCACTTCTCTGTGGAACTGCTGTGTCGGTCCGATCCCG	244
Db	59	LeuValLeuLeuLeuAlaAlaThrPheLeuThrPheLeuLeuValProValThrIlePro	78
QY	245	CGGGTCCGGTCCCTTCCACCGGTCGCCGATAACTTGTGGCTCGACGGCGCTCGGAC	304
Db	79	ArgValArgAlaPheHisArgValProHisAsnLeuValAlaSerThrAlaValSerAsp	98
QY	305	GAATAGTGGCAGCGCTCGCATGCCACCGAGCTCGGAGTGGCTGCACCGGCGCA	364
Db	99	GluLeuValAlaAlaLeuAlaWecProProSerLeuAlaSerGluLeuSerThrGlyArg	118
QY	365	CGTGGCTGCTGGGCGGAGCTGTGCCACGTGTGGATCTCTTCGACGCGCGAGCGTGT	424
Db	119	ArgArgLeuLeuGlyArgSerLeuCysHisValTrpIleSerPheAspAla-----	135
QY	425	CTGTGCTGCCCGCGCGCTCGGAACTGGCGGCCATCGCCCTGGCGCGCGCGGGCC	484
Db	136	LeuCysCysProAlaGlyLeuGlyAsnValAlaAlaIleAlaLeuGlyArgAspGlyAla	155
QY	485	ATCACACGGCACCTCAGCACAGCTGCGCACCGACCGCGCGCTGTCATCATC	544
Db	156	IleThrArgHisLeuGlnHisThrLeuArgThrArgSerArgAlaSerLeuLeuMetIle	175
QY	545	GGCTCGCGCGGTGCGCTCGCGCGCTCATCGCCCTCGCGCGCTCTTTGGCGCGGGC	604
Db	176	AlaLeuAlaArgValProSerAlaLeuIleAlaLeuAlaProLeuLeuPheGlyArgGly	195
QY	605	GAGGTGTCGAGCTCGGCTCCAGCGCTGCCAGTGGACCGGGAACTCTCTATGCCGCG	664
Db	196	GluValCysAspAlaArgLeuGlnArgCysGlnValSerArgGluProSerTyrAlaAla	215
QY	665	TTCTCCACCCCGCGCGCTCCACCTGCGCTTGGCGTGTGGCTTGTGTACCGGAG	724
Db	216	PheSerThrArgGlyAlaPheHisLeuProLeuGlyValValProPheValTyrArglys	235
QY	725	ATCTACGAGCGGCGCAAGTTTCGTTTCGGCGCGCGCGGAGAGCTGTGCTGCGCTGCG	784
Db	236	IleTyrGluAlaAlaLysPheArgPheGlyArgArgArgAlaValLeuProLeuPro	255
QY	785	GCCACCATGCAGTGCAGTCCAAAGTAAGAGACCACTCATGAGCTGAAGTGTGTTC	844
Db	256	AlaThrMetGln-----ValIleGluAlaProAspGluAlaGluValPhe	271
QY	845	ACGGCACATTGCAAGCAACGGTCTCTTCCAGGTGAGCGGACCTCTGGCGGGAGAG	904
Db	272	ThrAlaHisCysLysAlaThrValSerPheGlnValSerGlyAspSerTrpArgGluGln	291
QY	905	AAGGAGCGGAGCAGCCATGATCGTGGGAATCTCATTTGGCGGTGTTGTGCTGCTGG	964

Db	1	MetGluAlaAlaSerLeuSerValAlaThrAlaGlyValAlaLeuAlaLeuGlyProGlu	20
Qy	65	ACGAGCGAGCGAGCGGACCCCAAGCCCGAGAGGATACCTCGTTGACCCCGAGCGGC	124
Db	21	ThrSerSer-----GlyThrProSerProArgGlyIleLeuGlySerThrProSerGly	38
Qy	125	GCGGTCCTGCGGCGGAGCGCGCCCTCTCTGCTTCACGGTCTCGTGGTACGCTG	184
Db	39	AlaValLeuProGlyArgGlyProProPheSerValPheThrValValValThrLeu	58
Qy	185	CTAGTGTCTGATCGCTGCCACTTTCCTGTGGAACCTGCTGGTTCACCATCCG	244
Db	59	LeuValLeuLeuIleAlaAlaThrPheLeuTrpAsnLeuLeuValProValThrIlePro	78
Qy	245	CGGTCCTGCTTCCACCGGTCGCGCATAACTTGGTGGCTCGACGGCGCTCTCGGAC	304
Db	79	ArgValArgAlaPheHisArgValProHisAsnLeuValAlaSerThrAlaValSerAsp	98
Qy	305	GAACTAGTGGCAGCTGGCATGCCAGCCAGCTGGCGAGTGAGCTGTGACCGGGCGA	364
Db	99	GluLeuValAlaAlaLeuAlaMetProProSerLeuAlaSerGluLeuSerThrGlyArg	118
Qy	365	CCTGCTGCTGCGGCGGAGCTGTGCCACGTGTGGATCTCTTCAGCGCGGAGCCTGT	424
Db	119	ArgArgLeuLeuGlyArgSerLeuCySHisValTrpIleSerPheAspAla-----	135
Qy	425	CTGTGCTGCGCGCGCTCGGGAACGTGGCGGCATCGCCCTGGGCGCGACGGGCGC	484
Db	136	LeuCySHisProAlaGlyLeuGlyAsnValAlaAlaIleAlaLeuGlyArgAspGlyAla	155
Qy	485	ATCACAGCGCAGCTGACGACAGCTGCGCACCCGCGCGCTCGTTCATGATC	544
Db	156	IleThrArgHisLeuGlnHisThrLeuArgThrArgSerArgAlaSerLeuLeuMetIle	175
Qy	545	GCGCTGCGCGGCTCGCGCGCTCATCGCCCTCGCGCGCTGCTTTGGCGGGCGC	604
Db	176	AlaLeuAlaArgValProSerAlaLeuIleAlaLeuAlaProLeuLeuPheGlyArgGly	195
Qy	605	GAGGTGTGCGAGCTCGGCTCAGCGCTCGAGTGGAGCGGGAACCTCCTATCGCGC	664
Db	196	GluValCysAspAlaArgLeuGlnArgCysGlnValSerArgGluProSerTyrAlaAla	215
Qy	665	TTCTCCACCGCGCGCTTCACCTGCGCTGCGGTGCGGTGCTGCTACCGGAAG	724
Db	216	PheSerThrArgGlyAlaPheHisLeuProLeuGlyValValProPheValTyrArgLys	235
Qy	725	ATCTACGAGCGCGCAAGTTTCGTTTCGCGCGCGCGAGAGCTGCTGCGGTGCGG	784
Db	236	IleTyrGluAlaAlaLysPheArgPheGlyArgArgAlaValLeuProLeuPro	255
Qy	785	GCCACCATGAGGTGAGTCCAAAGTAAAGAGACCTGATGAGCTGAAGTGTGTTTC	844
Db	256	AlaThr-----SerLysValLysGluAlaProAspGluAlaGluValValPhe	271
Qy	845	ACGGCATTGCAAGCAACGCTGCTCTCCAGTGCAGCGGAGCTCCTGCGCGGAGCAG	904
Db	272	ThrAlaHisCysLysAlaThrValSerPheGlnValSerGlyAspSerTrpArgGluGln	291
Qy	905	AAGGAGAGGCGAGCAGCATGATGTTGGGAATTCATGATGGCGTGTGCTGCTGCTG	964
Db	292	LysGluArgArgAlaAlaMetMetValGlyIleLeuIleGlyValPheValLeuCysTrp	311
Qy	965	ATCCCTTCTTCTGACGAACTCATCGCCACTGTGCTGCTGCGAGCTGCGCCCATC	1024
Db	312	IleProPhePheLeuThrGluLeuLeuSerProLeuCysAlaCysSerLeuProPheIle	331
Qy	1025	TGCAAGAGCATATTTCTGTGCTGGCTACTCCAAATCTTCTTCAACCCCTGATTTAC	1084
Db	332	TrpLysSerIlePheLeuTrpLeuGlyTyrSerAsnSerPhePheAsnProLeuIleTyr	351
Qy	1085	ACAGCTTTTAAAGAACTACAATGCCTTCAAGAGCCTCTTTACTAAGCAGAGA	1141
Db	352	ThrAlaPheAsnLysAsnTyrAsnAsnAlaPheLysSerLeuPheThrLysGlnArg	370

RESULT 4

US-09-823-187-12

; Sequence 12, Application US/09823187

; Publication No. US2003009652A1

; GENERAL INFORMATION:

; APPLICANT: Burgess, Catherine

; APPLICANT: Gusev, Vladimir Y

; APPLICANT: Liu, Xiaohong

; APPLICANT: Majumder, Kumud

; APPLICANT: Padigaru, Muralidhar

; APPLICANT: Patturajan, Meera

; APPLICANT: Shimkets, Richard A

; APPLICANT: Spaderna, Steven K

; APPLICANT: Spytek, Kimberly

; APPLICANT: Taupier, Raymond J

; TITLE OF INVENTION: NOVEL PROTEINS AND NUCLEIC ACIDS ENCODING SAME

; FILE REFERENCE: 15966-745

; CURRENT APPLICATION NUMBER: US/09/823,187

; CURRENT FILING DATE: 2001-03-29

; PRIOR APPLICATION NUMBER: 60/193,339

; PRIOR FILING DATE: 2000-03-30

; PRIOR APPLICATION NUMBER: 60/193,205

; PRIOR FILING DATE: 2000-03-30

; PRIOR APPLICATION NUMBER: 60/195,343

; PRIOR FILING DATE: 2000-04-05

; PRIOR APPLICATION NUMBER: 60/195,088

; PRIOR FILING DATE: 2000-04-06

; PRIOR APPLICATION NUMBER: 60/195,005

; PRIOR FILING DATE: 2000-04-06

; PRIOR APPLICATION NUMBER: 60/195,792

; PRIOR FILING DATE: 2000-04-10

; PRIOR APPLICATION NUMBER: 60/196,556

; PRIOR FILING DATE: 2000-04-11

; PRIOR APPLICATION NUMBER: 60/197,081

; PRIOR FILING DATE: 2000-04-13

; PRIOR APPLICATION NUMBER: 60/197,525

; PRIOR FILING DATE: 2000-04-14

; PRIOR APPLICATION NUMBER: 60/197,087

; PRIOR FILING DATE: 2000-04-14

; NUMBER OF SEQ ID NOS: 103

; SOFTWARE: Patent in Ver. 2.1

; SEQ ID NO 12

; LENGTH: 370

; TYPE: PRT

; ORGANISM: Homo sapiens

US-09-823-187-12

Alignment Scores:

Pred. No.: 3,47e-107 Length: 370

Score: 1857.50 Matches: 369

Percent Similarity: 97.36% Conservative: 0

Best Local Similarity: 97.36% Mismatches: 1

Query Match: 86.23% Indels: 9

DB: 10 Gaps: 3

US-09-976-782-15 (1-1152) x US-09-823-187-12 (1-370)

Qy	5	ATGAGCGCGCTAGCCCTTTCAGTGGCCACCGCGCGGTTCCTTCCTCGGACCCGAG	64
Db	1	MetGluAlaAlaSerLeuSerValAlaThrAlaGlyValAlaLeuAlaLeuGlyProGlu	20
Qy	65	ACCAGAGCGGACCCCGGACCCCAAGCCCGAGAGGATACCTCGTTTCGACCCCGAGCGGC	124
Db	21	ThrSerSer-----GlyThrProSerProArgGlyIleLeuGlySerThrProSerGly	38
Qy	125	GCGTCTGCTGCGGCGGAGCGCGCCCTCTCTGCTTCACGGTCTCGTGGTACGCTG	184
Db	39	AlaValLeuProGlyArgGlyProProPheSerValPheThrValValValThrLeu	58
Qy	185	CTAGTGTCTGATCGCTGCCACTTTCCTGTGGAACCTGCTGGTTCACCATCCG	244
Db	59	LeuValLeuLeuIleAlaAlaThrPheLeuTrpAsnLeuLeuValProValThrIlePro	78

QY 245 CGGTCGCTGCTCCACCGCGTGCAGCCTTGGTGGCTCGACGCCGCTCGGAC 304
 Db |||||
 QY 79 ArgValArgAlaPheHisArgValProHisAsnLeuValAlaSerThrAlaValSerAsp 98
 Db |||||
 QY 305 GAACCTAGTGGCAGCGTGGCGATGCCACCGAGCTGGCGAGTGCAGTGCACCGGCGA 364
 Db |||||
 QY 99 GluLeuValAlaAlaLeuAlaMetProProSerLeuAlaSerGluLeuSerThrGlyArg 118
 Db |||||
 QY 365 CQTGGCTGCTGGCGCGAGCTGTGCCACGTGTGGATCTCTCGACCGCGAGCCTGT 424
 Db |||||
 QY 119 ArgArgLeuLeuGlyArgSerLeuCysHisValTrpIleSerPheAspAla 135
 Db |||||
 QY 425 CTGTCTGCTCCCGCGCGCTCGGCAACGTGGCGGCATCGCCCTGGCGCGCACGGGGC 484
 Db |||||
 QY 136 LeuCysCysProAlaGlyLeuGlyAsnValAlaAlaIleAlaLeuGlyArgAspGlyAla 155
 Db |||||
 QY 485 ATCACACGGCCTCGACGACACGCTGGCGACCCCGAGCGCGCTCGTGTCTATGATC 544
 Db |||||
 QY 156 IleThrArgHisLeuGlnHisThrLeuArgThrArgSerArgAlaSerLeuMetIle 175
 Db |||||
 QY 545 GCGCTCGCGCGGCTCGCGCTCATCGCCCTCGCGCGCTGCTTTGGCGGGC 604
 Db |||||
 QY 176 AlaLeuAlaArgValProSerAlaLeuAlaLeuAlaLeuAlaProLeuLeuPheGlyArgGly 195
 Db |||||
 QY 605 GAGTGTGCGACGCTCGGCTCCAGCGCTCCAGGTGAGCGCGGAACCTCTCTATCGCGC 664
 Db |||||
 QY 196 GluValCysAspAlaArgLeuGlnArgCysGlnValSerArgGluProSerTyrAlaAla 215
 Db |||||
 QY 665 TTCTCCACCGCGCGCTTCACCTGCGCGTGGCGTGGCGTGGCGTGGTGTCTACCGAAG 724
 Db |||||
 QY 216 PheSerThrArgGlyAlaPheHisLeuProLeuGlyValAlaProPheValTyrArgLys 235
 Db |||||
 QY 725 ATCTACGAGCGCGCAAGTTCTGTTTCGCGCGCGCGCGAGCTGCTCGCTGGTTC 784
 Db |||||
 QY 236 IleTyrGluAlaAlaLysPheArgPheGlyArgArgArgAlaValLeuProLeuPro 255
 Db |||||
 QY 785 GCCACCATGCAAGTGAGTCCCAAGTAAAGGAAGCACCTGATGAGCTGAAGTGGTGTTC 844
 Db |||||
 QY 256 AlaThrMetGln-----ValLysGluAlaProAspGluAlaGluValValPhe 271
 Db |||||
 QY 845 ACGGCATATCCAAAGCAACGGTGTCTTCCAGTGGAGCGGGACTCTCGCGGAGCAG 904
 Db |||||
 QY 272 ThrAlaHisCysLysAlaThrValSerPheGlnValSerGlyAspSerTyrArgGluGln 291
 Db |||||
 QY 905 AAGGAGGCGCGAGCAGCATCATGTTGGGAATCTCTGATGGCGTGTGCTGCTGCTGG 964
 Db |||||
 QY 292 LysGluArgAlaAlaMetMetValGlyIleLeuIleGlyValPheValLeuCysTyr 311
 Db |||||
 QY 965 ATCCCTTCTCTGACGGAACCTCATCAGCCCACTCTGTGCTGAGCTGCCCCCCTATC 1024
 Db |||||
 QY 312 IleProPheLeuThrGluLeuIleSerProLeuCysAlaCysSerLeuProIle 331
 Db |||||
 QY 1025 TGGAAAGCATATTTCTGTGGCTGGCTACTCAATTTCTTTCACCCCGCTGATTAC 1084
 Db |||||
 QY 332 TrpLysSerIlePheLeuTrpLeuGlyTyrSerAsnSerPhePheAsnProLeuIleTyr 351
 Db |||||
 QY 1085 ACAGCTTTTAAAGAACTCAACAATGCTTCAAGAGCCTCTTACTAAGCAGAGA 1141
 Db |||||
 QY 352 ThrAlaPheAsnLysAsnTyrAsnAsnAlaPheLysSerLeuPheThrLysGlnArg 370
 Db |||||

RESULT 5

US-10-311-671-8
 ; Sequence 8, Application US/10311671
 ; Publication No. US20040072996A1
 ; GENERAL INFORMATION:
 ; APPLICANT: INCYTE GENOMICS, INC.
 ; APPLICANT: TAL, Preeti G.
 ; APPLICANT: BAUGHN, Mariah R.
 ; APPLICANT: HAPALIA, April J. A.
 ; APPLICANT: NGUYEN, Daniel B.
 ; APPLICANT: GANDHI, Ameena R.
 ; APPLICANT: KALLICK, Deborah A.

; APPLICANT: GRIFPIN, Jennifer A.
 ; APPLICANT: YUE, Henry
 ; APPLICANT: KHAN, Farrah A.
 ; APPLICANT: ARVIZU, Chandra S.
 ; APPLICANT: LU, Dyung Aina M.
 ; APPLICANT: TRIBOULEY, Catherine M.
 ; APPLICANT: LU, Yan
 ; APPLICANT: CHAWLA, Narinder K.
 ; APPLICANT: GRAUL, Richard
 ; APPLICANT: YAO, Monique G.
 ; APPLICANT: YANG, Junming
 ; APPLICANT: RAMKUMAR, Jayalaxmi
 ; APPLICANT: AU-YOUNG, Janice K.
 ; APPLICANT: ELLIOTT, Vicki S.
 ; APPLICANT: HERNANDEZ, Roberto
 ; APPLICANT: WALSH, Roderick T.
 ; APPLICANT: BOROWSKY, Mark L.
 ; APPLICANT: THORNTON, Michael B.
 ; APPLICANT: HE, Ann
 ; TITLE OF INVENTION: G-PROTEIN COUPLED RECEPTORS
 ; FILE REFERENCE: PI-0131 USN
 ; CURRENT APPLICATION NUMBER: US/10/311,671
 ; CURRENT FILING DATE: 2002-12-16
 ; PRIOR APPLICATION NUMBER: PCT/US01/19275
 ; PRIOR FILING DATE: 2001-06-15
 ; PRIOR APPLICATION NUMBER: 60/212,483
 ; PRIOR FILING DATE: 2000-06-16
 ; PRIOR APPLICATION NUMBER: 60/213,954
 ; PRIOR FILING DATE: 2000-06-22
 ; PRIOR APPLICATION NUMBER: 60/215,209
 ; PRIOR FILING DATE: 2000-06-29
 ; PRIOR APPLICATION NUMBER: 60/216,595
 ; PRIOR FILING DATE: 2000-07-07
 ; PRIOR APPLICATION NUMBER: 60/218,936
 ; PRIOR FILING DATE: 2000-07-14
 ; PRIOR APPLICATION NUMBER: 60/219,154
 ; PRIOR FILING DATE: 2000-07-19
 ; PRIOR APPLICATION NUMBER: 60/220,141
 ; PRIOR FILING DATE: 2000-07-21
 ; NUMBER OF SEQ ID NOS: 35
 ; SOFTWARE: PERL Program
 ; SEQ ID NO 8
 ; LENGTH: 372
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 ; FEATURE:
 ; NAME/KEY: misc feature
 ; OTHER INFORMATION: Incyte ID No: 7474977CD1
 US-10-311-671-8

Alignment Scores:
 Pred. No.: 4,01e-107 Length: 372
 Score: 1856.50 Matches: 369
 Percent Similarity: 97.36% Conservative: 0
 Best Local Similarity: 97.36% Mismatches: 3
 Query Match: 86.19% Indels: 7
 DB: 12 Gaps: 3

US-09-976-782-15 (1-1152) x US-10-311-671-8 (1-372)

QY 5 ATGAGGCGCTAGCTTTCAGTGGCCACCGCGGGTTCCTTGGCTGGACCGCGAG 64
 Db |||||
 QY 1 MetGluAlaAlaSerLeuSerValAlaThrAlaGlyValAlaLeuAlaLeuGlyProGlu 20
 Db |||||
 QY 65 ACCAGCAGCGGACCCCGGACCCCAAGCCGAGAGGATCTCGGTTCGACCCCGAGCGGC 124
 Db |||||
 QY 21 ThrSerSer-----GlyThrProSerProArgGlyIleLeuGlySerThrProSerGly 38
 Db |||||
 QY 125 GCCGTCTCGCGGCGCGAGCGCGCCCTTCTCTGTCTTTCACGGTCTGTGTGTGAGCGCTG 184
 Db |||||
 QY 39 AlaValLeuProGlyArgGlyProProPheSerValPheThrValLeuValValThrLeu 58
 Db |||||
 QY 185 CTAGTCTCTGATCGCTGCCACTTTCCTGTGGAACTGCTGTGTTCCGTCACCATCCCG 244
 Db |||||

59	LeuValLeuLeuIleAlaAlaThrPheLeuTrpAsnLeuLeuValProValThrIlePro	78
245	CGGTCGCGCTTCACCGCTGGCGATACTTGGCTGCCTGCACGGCCGCTCTCGAC	304
79	ArgValArgAlaPheHisArgValProHisAsnLeuValAlaSerThrAlaValSerAsp	98
305	GAAC TAGTGGCAGCGTGGCGATGCCACCGAGCCTGGCGAGTGAGCTGTCGACCGGGCGA	364
99	GluLeuValAlaAlaLeuAlaMetProProSerLeuAlaSerGluLeuSerThrGlyArg	118
365	CGTGGCTGCTGGCCGGAGCCTGTCCACGCTGTGATCTCTTTCGACCGGAGCCTGT	424
119	ArgArgLeuLeuGlyArgSerLeuCysHisValTrpIleSerPheHisVal	135
425	CTGTGCTGCGCCGCGCTCGGGAACGTGGCGGCATCGCCCTGGCGCGCACGGGGCC	484
136	LeuCysCysProAlaGlyLeuGlyAsnValAlaAlaIleAlaLeuGlyArgAspGlyAla	155
485	ATCACACGGACACTGTCAGACACAGCTGCGGACCCGCGAGCGCGCTGTGTCATGATC	544
156	IleThrArgHisLeuGlnHisThrLeuArgThrArgSerArgAlaSerLeuLeuMetIle	175
545	CGCCTCGCCGGTGGCTGGCGCTCATCGCCCTCGCGCGCTGCTCTTGGCCGGGGC	604
176	AlaLeuThrArgValProSerAlaLeuIleAlaLeuAlaProLeuLeuPheGlyArgGly	195
605	GAGGTGTGCACGCTCGGCTCCACGCTGCCAGGTGAGCGGGAACCTCTATGCGCC	664
196	GluValCysAspAlaArgLeuGlnArgCysGlnValSerArgGluProSerTyAlaAla	215
665	TTCCTCACCGCGCGCTTCACACTGCGCTGGGTGGGTGGCTTGTCTACCGGAAG	724
216	PheSerThrArgGlyAlaPheHisLeuProLeuGlyValValProPheValTyArgLys	235
725	ATCTACGAGCGCCAGTTTCTTTCGCGCGCGCGGAGAGCTGCTGCTGCGCTGCGG	784
236	IleTyArgAlaAlaLysPheArgPheGlyArgArgArgAlaValLeuProLeuPro	255
785	GCCACCATGAGGTGAGTCCAAAGTAAAGAAAGACCTCATGAGGTGAAGTGGTGTTC	844
256	AlaThrMetGlnVal-----LysValLysGluAlaProAspGluAlaGluValPhe	273
845	ACGGCATTGCAAGCAACGGTGTCTTCACAGTGAGCGGGACTCCTCGCGGAGCAG	904
274	ThrAlaHisCysLysAlaThrValSerPheGlnValSerGlyAspSerTrpArgGluGln	293
905	AAGGAGAGCGGAGCAGCATGATGGTGGGAATCTGATGGCGGTGTTGTGCTGTGCTGG	964
294	LysGluArgArgAlaAlaMetMetValGlyIleLeuIleGlyValPheValLeuCysTrp	313
965	ATCCCTTCTTCAGCGGAACCTCATCGACCCACTCTGTGCTGCAGCCTGCCCCCATC	1024
314	IleProPheLeuThrGluLeuIleSerProLeuCysAlaCysSerLeuProProIle	333
1025	TGGAAGAAGCATATTCTGTGGCTGGCTACCTCAATCTTCTTCAACCCCTCATTTAC	1084
334	TrpLysSerIlePheLeuTrpLeuGlyTyrsSerAsnSerPhePheAsnProLeuIleTy	353
1085	ACAGCTTTTAAACAAGAACTACAACAATGCCTTCAAGAGCCTCTTACTTAAGCAGAGA	1141
254	ThrAlaPheAsnLysTrpAsnAlaPheLysSerLeuPheThrLysGlnArg	372

RESULT 6
US-09-954-342-30
; Sequence 30, Application US/09954342
; Publication No. US20030170838A1
; GENERAL INFORMATION:
; APPLICANT: MISHRA, VISHNU S.
; APPLICANT: SPYTEK, KIMBERLY ANN
; APPLICANT: TAUFIER, RAYMOND J.
; APPLICANT: CORNET, CORINE A.
; APPLICANT: VERNET, STEVEN D.
; APPLICANT: COLMAN, STEVEN D.

```

1  APPLICANT: GORMAN, LINDA
2  APPLICANT: TCHERNEV, VELIZAR T.
3  APPLICANT: MALYANKAR, URIEL M.
4  APPLICANT: SHENOY, SURESH
5  APPLICANT: PADIGARU, MURALIDHARA
6  APPLICANT: GERLACH, VALERIE L.
7  APPLICANT: MACDOUGALL, JOHN R.
8  APPLICANT: SMITHSON, GLENDA
9  APPLICANT: MILLET, ISABELLE
10 APPLICANT: PEYMAN, JOHN
11 APPLICANT: STONE, DAVID
12 APPLICANT: GUNTHER, BRIK
13 APPLICANT: ELLERMAN, KAREN
14 APPLICANT: LI, LI
15 APPLICANT: RASTELLI, LUCA
16 APPLICANT: ZERHUSEN, BRYAN
17
18 TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES AND POLYPEPTIDES ENCODED THEREBY
19
20 FILE REFERENCE: 21402-114
21
22 CURRENT APPLICATION NUMBER: US/09/954,342
23
24 CURRENT FILING DATE: 2001-09-17
25
26 PRIOR APPLICATION NUMBER: 60/233,382
27
28 PRIOR FILING DATE: 2000-09-18
29
30 PRIOR APPLICATION NUMBER: 60/240,498
31
32 PRIOR FILING DATE: 2000-10-13
33
34 PRIOR APPLICATION NUMBER: 60/260,284
35
36 PRIOR FILING DATE: 2001-01-08
37
38 PRIOR APPLICATION NUMBER: 60/260,973
39
40 PRIOR FILING DATE: 2001-01-11
41
42 PRIOR APPLICATION NUMBER: 60/264,794
43
44 PRIOR FILING DATE: 2001-01-29
45
46 PRIOR APPLICATION NUMBER: 60/238,398
47
48 PRIOR FILING DATE: 2000-10-06
49
50 PRIOR APPLICATION NUMBER: 60/232,675
51
52 PRIOR FILING DATE: 2000-09-15
53
54 PRIOR APPLICATION NUMBER: 60/274,862
55
56 PRIOR FILING DATE: 2001-03-09
57
58 PRIOR APPLICATION NUMBER: 60/233,801
59
60 PRIOR FILING DATE: 2000-09-19
61
62 PRIOR APPLICATION NUMBER: 60/232,676
63
64 PRIOR FILING DATE: 2000-09-15
65
66 PRIOR APPLICATION NUMBER: 60/233,960
67
68 PRIOR FILING DATE: 2000-09-20
69
70 PRIOR APPLICATION NUMBER: 60/233,402
71
72 PRIOR FILING DATE: 2000-09-18
73
74 PRIOR APPLICATION NUMBER: 60/233,521
75
76 PRIOR FILING DATE: 2000-09-19
77
78 PRIOR APPLICATION NUMBER: 60/233,522
79
80 PRIOR FILING DATE: 2000-09-19
81
82 PRIOR APPLICATION NUMBER: 60/232,679
83
84 PRIOR FILING DATE: 2000-09-15
85
86 NUMBER OF SEQ ID NOS: 104
87
88 SOFTWARE: PatentIn Ver. 2.1
89
90 SEQ ID NO 30
91
92 LENGTH: 380
93
94 TYPE: PRT
95
96 ORGANISM: Unknown Organism
97
98 FEATURE:
99
100 OTHER INFORMATION: Description of Unknown Organism: NOVX Polypeptide
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
841
842
843
844
845
846
847
848
```

Alignment Scores:	
Pred. No.:	7.94e-105
Score:	1819.50
Percent Similarity:	93.62%
Best Local Similarity:	93.62%
Query Match:	84.47%
DB:	10
Gaps:	5
Indels:	25
Mismatches:	0
Conservative:	0
Matches:	367
Length:	380

79 00-075-792-15 (1-1152) x 115-09-954-342-30 (1-380)

5 ATGGAGCCCGTAGCCCTTTCAGTGGCCACCGCCGCGTGGCCTTGGCCTGGGACCCGAG 64

```
QY 65 ACCAGCAGC-----GGACCCGGGACC 85
Db 19 ThrSerSerProAlaLeuProLeuProTrpAspProArgProAlaAlaGlyProGlyThr 38
QY 86 CCAAGCCGAGAGGATACCTGGTTCGACCCCGAGCGGCGCTGTCGCGGCGGAGG 145
Db 39 ProSerProArgGlyLeuGlySerThrProSerGlyAlaValLeuProGlyArgGly 58
QY 146 CGCCCTCTCTGCTCTTACCGTCTCGGTGGTGAGCGTCTAGTGTGTGTCGTCCTGCC 205
Db 59 ProProPheSerValPheThrValLeuValValThrLeuLeuValLeuLeuAla 78
QY 206 ACTTTCCTGTGAACCTGCTGCTTCGGTTCACCATCCCGCGGTCGTCGCTTCACCGCC 265
Db 79 ThrPheLeuTrpAsnLeuLeuValProValThrLeuProArgValArgAlaPheHisArg 98
QY 266 GTGCGCATACTTGGTGGCTCGAGCGGCGCTCTCGAGCAACTAGTGTGACGCTGGCG 325
Db 99 ValProHisAsnLeuValAlaSerThrAlaValSerAspGluLeuValAlaAlaLeuAla 118
QY 326 ATGCCACCGAGCCTGGCGAGTGTGCTGCGACCGGCGACGTCGGCTGTCGGCGGAGC 385
Db 119 MetProProSerLeuAlaSerGluLeuSerThrGlyArgArgLeuLeuGlyArg--- 137
QY 386 CTGTGCCACGTGTGGATCTCCTTTCGACCGCGGAGCCTGTGTGTGCTGCCCGCGGCTC 445
Db 138 -----HisValTrpIleSerPheAspAla-----LeuCysCysProAlaGlyLeu 152
QY 446 GGGAACTGGGGCCATCGCTTGGCGCGGCGGCGGCGGCGATCACACGGACCTGAGCAC 505
Db 153 GlyAsnValAlaAlaLeuLeuGlyArgAspGlyAlaIleThrArgHisLeuGlnHis 172
QY 506 ACGTGTGCGCACCGCAGCGCGCTGCTGTGCTATGATCGCGTGGCGCGGCGGCTGCG 565
Db 173 ThrLeuArgThrArgSerArgAlaSerLeuLeuMetIleAlaLeuAlaArgValProSer 192
QY 566 GCGTCTATCGCCCTCGCGCGCTGCTTCTTGGCGGCGGAGGTGTGCGAGCTCGGCTC 625
Db 193 AlaLeuIleAlaLeuAlaProLeuLeuPheGlyArgGlyGluValCysAspAlaArgLeu 212
QY 626 CAGCGTGCAGAGTGTGAGCGGGAACCTCTATGCGCGCTTCTCCACCGCGGCGCTTC 685
Db 213 GlnArgCysGlnValSerArgGluProSerTyrAlaAlaPheSerThrArgGlyAlaPhe 232
QY 686 CACCTGCGCTGTGCGTGTGCTGCTACCGGAAGATCTACGAGCGGCGGCGCACTTT 745
Db 233 HisLeuProLeuGlyValValProPheValTyrArgLysIleTyrGluAlaAlaLysPhe 252
QY 746 CGTTTCGGCGCGCGGAGAGCTGTGCTGCGGTGCGCGGCGGCGGCGGCGGAGGTGCC 805
Db 253 ArgPheGlyArgArgArgAlaValLeuProLeuProAlaThrMetGln----- 269
QY 806 AAGTAAGGAGACCTGTATGAGCTGAGTGTGTCACGGCACATTCGAAGCAAGC 865
Db 270 ---ValLysGluAlaProAspGluAlaGluValValPheThrAlaHisCysLysAlaThr 288
QY 866 GTGTCTTCAGGAGAGCGGAGCTCTCGCGGAGCAGAGGAGGAGCGGCGGCGGCGG 925
Db 289 ValSerPheGlnValSerGlyAspSerTrpArgGluGlnLysGluArgAlaAlaMet 308
QY 926 ATGGTGGGAATCTGATGGCGTGTGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 985
Db 309 MetValGlyIleLeuIleGlyValPheValLeuCysTrpIleProPhePheLeuThrGlu 328
QY 986 CTATCAGCCCACTCTGTGCTGCGAGCTGCGCCCACTGGAAGCAATATTTCTGTGG 1045
Db 329 LeuIleSerProLeuCysAlaCysSerLeuProProIleTrpLysSerIlePheLeuTrp 348
QY 1046 CTGGCTACTCCAAATCTTCTTCAACCCCTGATTTTACACAGCTTTTAAACAAGAACTAC 1105
Db 349 LeuGlyTyrSerAsnSerPhePheAsnProLeuIleTyrThrAlaPheAsnLysAsnTyr 368
```

```
QY 1106 AACATGCTCTCAAGAGCCTCTTTACTAAGCAGAGA 1141
Db 369 AsnAsnAlaPheLysSerLeuPheThrLysGlnArg 380
RESULT 7
US-09-823-187-56
; Sequence 56, Application US/09823187
; Publication No. US20030096952A1
; GENERAL INFORMATION:
; APPLICANT: Burgess, Catherine
; APPLICANT: Gusev, Vladimir Y
; APPLICANT: Liu, Xiaohong
; APPLICANT: Majumder, Kumud
; APPLICANT: Patturajan, Muralidhar
; APPLICANT: Shinkets, Richard A
; APPLICANT: Spaderna, Steven K
; APPLICANT: Spytek, Kimberly
; APPLICANT: Taupier, Raymond J
; TITLE OF INVENTION: NOVEL PROTEINS AND NUCLEIC ACIDS ENCODING SAME
; FILE REFERENCE: 15966-745
; CURRENT APPLICATION NUMBER: US/09/823,187
; CURRENT FILING DATE: 2001-03-29
; PRIOR APPLICATION NUMBER: 60/193,339
; PRIOR FILING DATE: 2000-03-30
; PRIOR APPLICATION NUMBER: 60/193,205
; PRIOR FILING DATE: 2000-03-30
; PRIOR APPLICATION NUMBER: 60/195,343
; PRIOR FILING DATE: 2000-04-05
; PRIOR APPLICATION NUMBER: 60/195,088
; PRIOR FILING DATE: 2000-04-06
; PRIOR APPLICATION NUMBER: 60/195,005
; PRIOR FILING DATE: 2000-04-06
; PRIOR APPLICATION NUMBER: 60/195,792
; PRIOR FILING DATE: 2000-04-10
; PRIOR APPLICATION NUMBER: 60/196,556
; PRIOR FILING DATE: 2000-04-11
; PRIOR APPLICATION NUMBER: 60/197,081
; PRIOR FILING DATE: 2000-04-13
; PRIOR APPLICATION NUMBER: 60/197,525
; PRIOR FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/197,087
; NUMBER OF SEQ ID NOS: 103
; SOFTWARE: PatentIn ver. 2.1
; SEQ ID NO 56
; LENGTH: 370
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-823-187-56
Alignment Scores:
Pred. No.: 4,04e-82 Length: 370
Score: 1453.50 Matches: 296
Percent Similarity: 83.38% Conservative: 20
Best Local Similarity: 78.10% Mismatches: 54
Query Match: 67.48% Indels: 9
DB: 10 Gaps: 3
US-09-976-782-15 (1-1152) x US-09-823-187-56 (1-370)
QY 5 ATGAGCGCCCTAGCCTTTTCAGTGGCCACCGCGCGCTTGCCTTGCCTGGAGCCGAG 64
Db 1 MetGluValSerAsnLeuSerGlyAlaThrProGlyIleAlaPheProProGlyProGlu 20
QY 65 ACCAGCAGCGGACCCCGGACCCCAAGCCCGAGAGGATACCTCGGTTTCGACCCCGAGCGGC 124
Db 21 SerCysSer-----AspSerProSerSerGlyArgSerMetGlySerThrProGlyGly 38
QY 125 GCGCTCTCGCGGCGGCGGCGGCGGCTTCTCTGTCTTTCACGGTCTCTGTGTGTGAGCGCTG 184
Db 39 LeuIleLeuSerGlyArgGluProProPheSerAlaPheThrValLeuValThrLeu 58
```


Db 119 ArgTrpGlnLeuGlyArgSerLeuCyshisValTrpIleSerPheAspVal----- 135
Qy 425 CTGTGCTGCGCGCGCTCGGAGACTGCGGCATCGCGCCCTCGCGCGCGAGCGGCGCC 484
Db 136 LeuCyshisThrAlaSerIleTrpAsnValAlaAlaIleAlaLeuAspArgTyrTrpThr 155
Qy 485 ATCACAGCGCACCTGCAGCACACGCTGCGCACCGCGCGCGCGCGCTCGCTCATGATC 544
Db 156 IleThrArgHisLeuGlnTyrThrLeuArgThrArgArgAlaSerAlaLeuMetIle 175
Qy 545 GCGTCTCGCGCGCGCTCGCGCTCATGCGCGCTCGCGCGCGCTCGCTCGCTCGCGCGCG 604
Db 176 AlaIleThrTrpAlaLeuSerAlaLeuIleAlaLeuAlaProLeuLeuPheGlyTrpGly 195
Qy 605 GAGTGTGCGACGCTCGCGCTCGCGCTCGCGCTCGCGCTCGCGCTCGCGCTCGCGCTCG 664
Db 196 GluAlaTyrAspAlaArgLeuGlnArgCysGlnValSerGlnGluProSerTyrAlaVal 215
Qy 665 TTCTCCACCGCGCGCGCTCGCGCTCGCGCTCGCGCTCGCGCTCGCGCTCGCGCTCGCG 724
Db 216 PheSerThrCysGlyAlaPheTyrValProLeuAlaValValLeuPheValTyrTrpLys 235
Qy 725 ATTCAGAGCGCGCAAGTTTGTGTTTGTGTTTGTGTTTGTGTTTGTGTTTGTGTTTGTG 784
Db 236 IleTyrIysAlaAlaLysPheArgPheGlyArgArgAlaValValProLeuPro 255
Qy 785 GCACCATGAGGTGAGGTCCAGGTAAAGAGACACCTGATGAGGTGAGGTGAGGTGAG 844
Db 256 AlaThrThrGln-----AlaLysGluAlaProGlnGluSerGluThrValPhe 271
Qy 845 ACAGCATTTGCAAGCAAGTGTCTCTCCAGGTGAGCGGGGACTCTCGCGGGGAGCAG 904
Db 272 ThrAlaArgCysArgAlaThrValAlaPheGlnThrSerGlyAspSerTrpArgGluGln 291
Qy 905 AAGAGAGCGGAGCAGCATGATGTTGGGAATTCGATGTCGCGGTGTTGCTGTGCTGG 964
Db 292 LysGluLysArgAlaAlaMetMetValGlyIleLeuIleGlyValPheValLeuCysTrp 311
Qy 965 ATCCCTTCTCTCGCGGACTCATAGCCCACTGTCGCTGCGCGCGCGCGCGCGCGCG 1024
Db 312 IleProPhePheLeuThrGluLeuValSerProLeuCysAlaCysSerLeuProIle 331
Qy 1025 TGGAAAAGCATATTTCTGTGCTTGGCTACTCTCAATTTCTTCTCAACCCCGCTGATTAC 1084
Db 332 TrpLysSerIlePheLeuTrpLeuGlyTyrSerAsnSerPhePheAsnProLeuIleTyr 351
Qy 1085 ACAGCTTTTAAAGAACTCAACAATGCTTCAAGAGCTCTTTTAAAGAGAGA 1141
Db 352 ThrAlaPheAsnLysAsnTyrAsnAsnAlaPheLysSerLeuPheThrLysGlnArg 370

RESULT 9
US-09-954-342-56
; Sequence 56, Application US/09954342
; Publication No. US20030170838A1
; GENERAL INFORMATION:
; APPLICANT: MISHRA, VISHNU S.
; APPLICANT: SPYTEK, KIMBERLY ANN
; APPLICANT: TAUPIER, RAYMOND J.
; APPLICANT: VERNET, CORINE A.
; APPLICANT: COLMAN, STEVEN D.
; APPLICANT: GORMAN, LINDA
; APPLICANT: TCHERNEV, VELIZAR T.
; APPLICANT: MALYANKAR, URIEL M.
; APPLICANT: SHENOY, SURESH
; APPLICANT: PADIGARU, MURALIDHARA
; APPLICANT: GERLACH, VALERIE L.
; APPLICANT: MACDOUGALL, JOHN R.
; APPLICANT: SMITHSON, GLENDA
; APPLICANT: MILLER, ISABELLE
; APPLICANT: PEYMAN, JOHN
; APPLICANT: STONE, DAVID
; APPLICANT: GUNTHER, ERIK

; APPLICANT: ELLERMAN, KAREN
; APPLICANT: LI, LI
; APPLICANT: RASTELLI, LUCA
; APPLICANT: ZERHUSEN, BRYAN
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES AND POLYPEPTIDES ENCODED THEREBY
; FILE REFERENCE: 21402-114
; CURRENT APPLICATION NUMBER: US/09/954,342
; PRIOR FILING DATE: 2001-09-17
; PRIOR APPLICATION NUMBER: 60/233,382
; PRIOR FILING DATE: 2000-09-18
; PRIOR APPLICATION NUMBER: 60/240,498
; PRIOR FILING DATE: 2000-10-13
; PRIOR APPLICATION NUMBER: 60/260,284
; PRIOR FILING DATE: 2001-01-08
; PRIOR APPLICATION NUMBER: 60/260,973
; PRIOR FILING DATE: 2001-01-11
; PRIOR APPLICATION NUMBER: 60/264,794
; PRIOR FILING DATE: 2001-01-29
; PRIOR APPLICATION NUMBER: 60/238,398
; PRIOR FILING DATE: 2000-10-06
; PRIOR APPLICATION NUMBER: 60/232,675
; PRIOR FILING DATE: 2000-09-15
; PRIOR APPLICATION NUMBER: 60/274,862
; PRIOR FILING DATE: 2001-03-09
; PRIOR APPLICATION NUMBER: 60/233,801
; PRIOR FILING DATE: 2000-09-19
; PRIOR APPLICATION NUMBER: 60/232,676
; PRIOR FILING DATE: 2000-09-15
; PRIOR APPLICATION NUMBER: 60/233,960
; PRIOR FILING DATE: 2000-09-20
; PRIOR APPLICATION NUMBER: 60/233,402
; PRIOR FILING DATE: 2000-09-18
; PRIOR APPLICATION NUMBER: 60/233,521
; PRIOR FILING DATE: 2000-09-19
; PRIOR APPLICATION NUMBER: 60/233,522
; PRIOR FILING DATE: 2000-09-19
; PRIOR APPLICATION NUMBER: 60/232,679
; NUMBER OF SEQ ID NOS: 104
; SOFTWARE: Patent in ver. 2.1
; SEQ ID NO 56
; LENGTH: 370
; TYPE: PRT
; ORGANISM: Rattus norvegicus
US-09-954-342-56

Alignment Scores:
Pred. No.: 4,04e-82 Length: 370
Score: 1453.50 Matches: 296
Percent Similarity: 83.38% Conservative: 20
Best Local Similarity: 78.10% Mismatches: 54
Query Match: 67.48% Indels: 9
DB: 10 Gaps: 3

US-09-976-782-15 (1-1152) x US-09-954-342-56 (1-370)

Qy 5 ATGAGCGCCCTAGCCTTTCAGTGGCCACCGCGCGGCTTGCCTTGCCTTGGGACCGGAG 64
Db 1 MetGluValSerAsnLeuSerGlyAlaThrProGlyIleAlaPheProGlyProGlu 20
Qy 65 ACCAGCAGCGGACCGCGGACCCCAAGCCCGAGAGGATACCTCGGTTCCGACCCCGAGCGGC 124
Db 21 SerCysSer-----AspSerProSerSerGlyArgSerMetGlySerThrProGlyGly 38
Qy 125 GCAGTCTCTCGCGCGCGGAGCGCGCTTCTCTGTCTTTCACGGTCTCTGGTGGTGCAGCTG 184
Db 39 LeuIleLeuSerGlyArgGluProProPheSerAlaPheThrValLeuValThrLeu 58
Qy 185 CTAGTGTCTGTGATCGCTGCCACTTTCCTGTGGAACCTTCTGTGGAACCTTCTGTGGAACCT 244
Db 59 LeuValLeuLeuIleAlaAlaThrPheLeuTrpAsnLeuLeuValLeuValThrLeu 78
Qy 245 CGGGTCCGTGCCTTCCACCGCGTGGCCGATACCTTGTGTGGCTCGACGGCGGCTCTCGGAC 304

```
Db 79 ArgValArgAlaPheHisArgValProHisnLeuValAlaSerThrAlaValSerAsp 98
QY 305 GAACATAGTGGCAGCGCTCGCGATGCCACCGAGCTGGCGAGTGAGCTGTGCACCGGGCGA 364
Db 99 ValLeuValAlaAlaLeuValMetProLeuSerLeuValSerGluLeuSerAlaGlyArg 118
QY 365 CGTGGCTGCTGGCGGAGCTGGCCACGTGGATCTCTTCGACGCCGAGCCTGT 424
Db 119 ArgTrpGlnLeuGlyArgSerLeuCysHisValTrpIleSerPheAspVal----- 135
QY 425 CTGTGCTGCCCGCGCTCGGGAACGTGGCGGCATCGCTGGCGCGCGACGGGGCC 484
Db 136 LeuCysCysThrAlaSerIleTrpAsnValAlaAlaIleAlaLeuAspArgTrpThr 155
QY 485 ATCACAGCGCATCTGCAGCACACGTGGCGACCGCGCGCGCTGTGTCTCATGATC 544
Db 156 IleThrArgHisLeuGlnTyrThrLeuArgThrArgArgAlaSerAlaLeuMetIle 175
QY 545 GCGCTCGCCGGTCCGCTCGCGCTCATCGCCCTCGCGCGCTCTTTGGCGGGGC 604
Db 176 AlaIleThrTrpAlaLeuSerAlaLeuIleAlaLeuAlaProLeuLeuPheGlyTrpGly 195
QY 605 GAGGTGTGCGAGCTCGCTCGAGCGCTGCCAGGTGCCCGGGAACCCCTCTATCGGCC 664
Db 196 GluAlaTyrAspAlaArgLeuGlnArgCysGlnValSerGlnGluProSerTyrAlaVal 215
QY 665 TTCTCCACCGCGGCGCTTCCACCTCGCGCTGGCGGTGGCGGTGGTGTCTACCGGAAG 724
Db 216 PheSerThrCysGlyAlaPheTyrValProLeuAlaValValLeuPheValTyrTrpLys 235
QY 725 ATCTACGAGCGCGCAAGTTCTGTTTCGCGCGCGCGAGAGCTGTGCTGCGTGGCG 784
Db 236 IleTyrLysAlaAlaLysPheArgPheGlyArgArgAlaValValProLeuPro 255
QY 785 GCCACCATGTCAGGTGAGGTCCAAAGGTAAAGAGACACCTGATGAGCTGAAGTGTTC 844
Db 256 AlaThrThrGln-----AlaLysGluAlaProGlnGluSerGluThrValPhe 271
QY 845 AGCGACATGTCAAAGCAACGCTGCTTCACAGGTGAGCGGGGACTCTCGCGGGAGCAG 904
Db 272 ThrAlaArgCysArgAlaThrValAlaPheGlnThrSerGlyAspSerTrpArgGluGln 291
QY 905 AAGGAGCGCAGCAGCATGATGTGGGAATCTGATTGGCGTGTGTGCTGTGCTGG 964
Db 292 LysGluLysArgAlaAlaMetMetValGlyIleLeuIleGlyValPheValLeuCysTrp 311
QY 965 ATCCCTCTCTCTGACGGAATCATFACGCCCACTGTGCTGTCGAGCCTGCCCCCATC 1024
Db 312 IleProPhePheLeuThrGluLeuValSerProLeuCysAlaCysSerLeuProPhe 331
QY 1025 TGGAAAAGCATATTTCTGTGGCTGTGCTACTCGAATCTTTCTCAACCCCTGATTAC 1084
Db 332 TrpLysSerIlePheLeuTrpLeuGlyTyrSerAsnSerPhePheAsnProLeuIleTyr 351
QY 1085 ACAGCTTTTAAACAAGACTACACATGCTTCAAGAGCTCTTACTAAGCAGAGA 1141
Db 352 ThrAlaPheAsnLysAsnLysTrpAsnAsnAlaPheLysSerLeuPheThrLysGlnArg 370
```

RESULT 10

```
US-09-976-782-101
; Sequence 101, Application US/09976782
; Publication No. US20030190715A1
; GENERAL INFORMATION:
; APPLICANT: Grosse et al
; TITLE OF INVENTION: No. US20030190715A1e1 Proteins and Nucleic Acids Encoding Same
; FILE REFERENCE: 21402-157
; CURRENT APPLICATION NUMBER: US/09/976,782
; CURRENT FILING DATE: 2001-10-12
; PRIOR APPLICATION NUMBER: 60/240,113
; PRIOR FILING DATE: 2000-10-12
; PRIOR APPLICATION NUMBER: 60/240,662
; PRIOR FILING DATE: 2000-10-16
```

```
; PRIOR APPLICATION NUMBER: 60/240,732
; PRIOR FILING DATE: 2000-10-16
; PRIOR APPLICATION NUMBER: 60/240,625
; PRIOR FILING DATE: 2000-10-16
; PRIOR APPLICATION NUMBER: 60/240,703
; PRIOR FILING DATE: 2000-10-16
; PRIOR APPLICATION NUMBER: 60/241,190
; PRIOR FILING DATE: 2000-10-16
; PRIOR APPLICATION NUMBER: 60/240,637
; PRIOR FILING DATE: 2000-10-16
; PRIOR APPLICATION NUMBER: 60/240,669
; PRIOR FILING DATE: 2000-10-16
; PRIOR APPLICATION NUMBER: 60/262,455
; PRIOR FILING DATE: 2001-01-18
; PRIOR APPLICATION NUMBER: 60/240,648
; PRIOR FILING DATE: 2000-10-16
; NUMBER OF SEQ ID NOS: 127
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 101
; LENGTH: 370
; TYPE: PRT
; ORGANISM: Rattus norvegicus
US-09-976-782-101

Alignment Scores:
Pred. No.: 4,04e-82 Length: 370
Score: 1453.50 Matches: 296
Percent Similarity: 83.38% Conservative: 20
Best Local Similarity: 78.10% Mismatches: 54
Query Match: 67.48% Indels: 9
DB: 10 Gaps: 3
```

US-09-976-782-15 (1-1152) x US-09-976-782-101 (1-370)

```
QY 5 ATGAGGCGCGTAGCTTTTCAGTGCCACCGCGCGGTGTCCTTGCCTGGAGCCCGAG 64
Db 1 MetGluValSerAsnLeuSerGlyAlaThrProGlyIleAlaPheProGlyProGlu 20
QY 65 ACCAGCAGCGGACCCCGGACCCCAAGCCGAGAGGGATCTCGTTCCAGCCCGCGCGC 124
Db 21 SerCysSer-----AspSerProSerSerGlyArgSerMetGlySerThrProGlyGly 38
QY 125 GCGGTCTCTCGCGCGCGAGCGCCCTTCTCTGTCTTCCACGGTCTGTGTGTGACGCTG 184
Db 39 LeuIleLeuSerGlyArgGluProPheSerAlaPheThrValLeuValThrLeu 58
QY 185 CTAGTGTCTGATCGCTGCCACTTCTGTGGAACTCTGTTCGGTTCGGTTCACCATCCCG 244
Db 59 LeuValLeuLeuIleAlaAlaThrPheLeuTrpAsnLeuLeuValLeuThrIleLeu 78
QY 245 CCGGTCTCGTCTTCCACCGCTCGCGCAATACTTGGTGGCTCGACGGCGCTCTCGGAC 304
Db 79 ArgValArgAlaPheHisArgValProHisAsnLeuValAlaSerThrAlaValSerAsp 98
QY 305 GAACATAGTGGCAGCGCTGGCGATGCCACCGAGCCTGGCGAGTGAGTGTGCACCGGGCGA 364
Db 99 ValLeuValAlaAlaLeuValMetProLeuSerLeuValSerGluLeuSerAlaGlyArg 118
QY 365 CGTGGCTGCTGGCGGAGCCTGTGCCACGTGGATCTCTTCGACGCCGAGCCTGT 424
Db 119 ArgTrpGlnLeuGlyArgSerLeuCysHisValTrpIleSerPheAspVal----- 135
QY 425 CTGTGCTGCCCGCGCTCGGGAACGTGGCGGCATCGCTGGCGCGCGACGGGGCC 484
Db 136 LeuCysCysThrAlaSerIleTrpAsnValAlaAlaIleAlaLeuAspArgTrpThr 155
QY 485 ATCACAGCGCATCTGCAGCACACGTGGCGACCGCGCGCGCTGTGTCTCATGATC 544
Db 156 IleThrArgHisLeuGlnTyrThrLeuArgThrArgArgAlaSerAlaLeuMetIle 175
QY 545 GCGCTCGCCGGTCCGCTCGCGCTCATCGCCCTCGCGCGCTCTTTGGCGGGGC 604
Db 176 AlaIleThrTrpAlaLeuSerAlaLeuIleAlaLeuAlaProLeuLeuPheGlyTrpGly 195
```



```
QY 605 GAGGTGTGCGACGCTCGGCTCCAGCGCTGCCAGGTGAGCCGGGAACCTCCCTATCCCGCC 664
Db 196 GluAlaTyrAspAlaArgLeuGlnArgCysGlnValSerGlnLupProSerTyrAlaVal 215
QY 665 TTCTCCACCGCGCGCTCCACCTGCGCTTGGCGTGGTGGTGGTGGTGGTGGTGGTGGTGG 724
Db 216 PheSerThrCysGlyAlaPheTyrValProLeuAlaValValLeuPheValTyrTrpLys 235
QY 725 ATCTACGAGGCGCGCAAGTTTGGTTTGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 784
Db 236 IleTyrLysAlaAlaLysPheArgPheGlyArgArgArgAlaValValProLeuPro 255
QY 785 GCCACCATGCGAGTGCAGGTCACAGCTAAAGGAGACCTGATGAGCTGAAGTGTGTTC 844
Db 256 AlaThrThrGln-----AlaLysGluAlaProGlnLupSerGluThrValPhe 271
QY 845 ACGGCACATGTCAGAACAGCGTGTCTCCAGGTGAGCGGGGACTCTCGCGGGGAGCAG 904
Db 272 ThrAlaArgCysArgAlaThrValAlaPheGlnThrSerGlyAspSerTyrArgGluGln 291
QY 905 AAGGAGGCGGAGCAGCATGATGCTGGGAAATCTGATGGGGTGTGTGTGTGTGTGTGTG 964
Db 292 LysGluLysArgAlaAlaMetMetValGlyIleLeuIleGlyValPheValLeuCysTrp 311
QY 965 ATCCCGCTTCTTCTGACGGAATCATACAGCCACTCTGTGCTGCGAGCTGCCCGCCATC 1024
Db 312 IleProPhePheLeuThrGluLeuValSerProLeuCysAlaCysSerLeuProPhe 331
QY 1025 TGGAAAGCATATTCTGTGCTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 1084
Db 332 TrpLysSerIlePheLeuThrLeuGlyTyrSerAsnSerPhePheAsnProLeuIleTyr 351
QY 1085 ACAGCTTTTACAGAACTCAACAACTGCTTCAAGAGCTTCTTACTAAGCAGAGA 1141
Db 352 ThrAlaPheAsnLysAsnTyrAsnAsnAlaPheLysSerLeuPheThrLysGlnArg 370

RESULT 11
US-09-823-187-58
; Sequence 58: Application US/09823187
; Publication No. US20030096952A1
; GENERAL INFORMATION:
; APPLICANT: Burgess, Catherine
; APPLICANT: Gusev, Vladimir Y
; APPLICANT: Liu, Xiaohong
; APPLICANT: Majumder, Kumud
; APPLICANT: Padigaru, Muralidhar
; APPLICANT: Patturajan, Meera
; APPLICANT: Shimkets, Richard A
; APPLICANT: Spaderna, Steven K
; APPLICANT: Spyttek, Kimberly
; APPLICANT: Taupier, Raymond J
; TITLE OF INVENTION: NOVEL PROTEINS AND NUCLEIC ACIDS ENCODING SAME
; FILE REFERENCE: 15966-745
; CURRENT APPLICATION NUMBER: US/09/823,187
; CURRENT FILING DATE: 2001-03-29
; PRIOR APPLICATION NUMBER: 60/193,339
; PRIOR FILING DATE: 2000-03-30
; PRIOR APPLICATION NUMBER: 60/193,205
; PRIOR FILING DATE: 2000-03-30
; PRIOR APPLICATION NUMBER: 60/195,343
; PRIOR FILING DATE: 2000-04-05
; PRIOR APPLICATION NUMBER: 60/195,088
; PRIOR FILING DATE: 2000-04-06
; PRIOR APPLICATION NUMBER: 60/195,005
; PRIOR FILING DATE: 2000-04-06
; PRIOR APPLICATION NUMBER: 60/195,792
; PRIOR FILING DATE: 2000-04-10
; PRIOR APPLICATION NUMBER: 60/196,556
; PRIOR FILING DATE: 2000-04-11
; PRIOR APPLICATION NUMBER: 60/197,081
; PRIOR FILING DATE: 2000-04-13
; PRIOR APPLICATION NUMBER: 60/197,525
```

```
; PRIOR FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/197,087
; PRIOR FILING DATE: 2000-04-14
; NUMBER OF SEQ ID NOS: 103
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 58
; LENGTH: 370
; TYPE: PRT
; ORGANISM: Mus musculus
US-09-823-187-58

Alignment Scores:
Pred. No.: 4,66e-82 Length: 370
Score: 1452.50 Matches: 298
Percent Similarity: 83.64% Conservative: 19
Best local Similarity: 78.63% Mismatches: 53
Query Match: 67.43% Indels: 9
DB: 10 Gaps: 3
```

```
US-09-976-782-15 (1-1152) x US-09-823-187-58 (1-370)
QY 5 ATGAGAGCGCGTAGCCTTTTCAGTGGCCACCGCGCGCTTGGCCCTTGCCTCGGACCCGAG 64
Db 1 MetGluValSerAsnLeuSerGlyAlaThrProGlyLeuAlaPheProProGlyProGlu 20
QY 65 ACCAGCAGCGGACCCCGGAGCCCAAGCCGAGAGGAGGATCTCGGTTGACCCCGAGCGGC 124
Db 21 SerCysSer-----AspSerProSerSerGlyArgSerMetGlySerThrProGlyGly 38
QY 125 GCCGTCTCGCGCGCGCGCGCGCTTCTCTGCTTCTACGGTCTCGGTGCTGACGCTG 184
Db 39 LeuIleLeuProGlyArgGluProPheSerAlaPheThrValLeuValValThrLeu 58
QY 185 CTAGTGTGTGATCGCTGCGCACCTTCTGTGGAACCTGCTGGTCCGGTCCAGTATCCG 244
Db 59 LeuValLeuLeuIleAlaAlaThrPheLeuTrpAsnLeuLeuValLeuValThrLeu 78
QY 245 CGGGTCTGCTGCTTCCACCGCGTCCGCACTTGGTGGCTCGAGCGCGCTCTCGGAC 304
Db 79 ArgValArgAlaPheHisArgValProHisAsnLeuValAlaSerThrAlaValSerAsp 98
QY 305 GAAGTAGTGGCAGCGCTGGCGATGCCACCGAGCTGGCGAGTGTCTCGACCGCGGCGA 364
Db 99 ValLeuValAlaValLeuValMetProLeuSerLeuValSerGluLeuSerAlaGlyArg 118
QY 365 CGTGGCTGTGGCGCGGAGCCTGTGCGACCTGTGTGGATCTCTCTCGACCGCGGACCTGT 424
Db 119 ArgTrpGlnLeuGlyArgSerLeuCysHisValTrpIleSerPheAspVal----- 135
QY 425 CTGTGCTGCCCGCGCGCTCGGGAACGTGGCGGCCCATCGCCCTGGCGCGGCGGCGCC 484
Db 136 LeuCysThrAlaSerIleTrpAsnValAlaAlaIleAlaLeuAspArgTyrTrpThr 155
QY 485 ATCACAGCGCACCTGCAGCACACCTGCGCACCGCGCGCGCTCGTGTCTCATGATC 544
Db 156 IleThrArgHisLeuGlnTyrThrLeuArgThrArgSerArgAlaSerAlaLeuMetIle 175
QY 545 GCGCTCGCGCGGCGCGCTCATCGCCTCGCGCGCTGTCTCTTTGCGCGCGGCG 604
Db 176 AlaIleThrTrpAlaLeuSerAlaLeuIleAlaLeuAlaProLeuLeuPheGlyTrpGly 195
QY 605 GAGGTGTGCGAGCGCTCGGCTCCAGCGTGGCAGGTGAGCGCGGAGCCCTCTATGCGGCC 664
Db 196 GluAlaTyrAspAlaArgLeuLeuArgCysGlnValSerGlnLupProSerTyrAlaVal 215
QY 665 TTCTCCACCGCGCGCTTCCACCTGCGCTTGGCGTGGTGGTGGTGGTGGTGGTGGTGGTGG 724
Db 216 PheSerThrCysGlyAlaPheTyrLeuProLeuAlaValValLeuPheValTyrTrpLys 235
QY 725 ATCTACGAGGCGCGCAAGTTTGGTTTGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 784
Db 236 IleTyrLysAlaAlaLysPheArgPheGlyArgArgArgAlaValValProLeuPro 255
```

```
QY 785 GCCACATGCGAGTCCAGGTCCAAAGTAAAGGAAGCACCTGATGAGGTGAAGTGTGTTTC 844
Db ||||| ||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
256 AlaThrThrGln-----AlaLysGluAlaProGluSerGluMetValPhe 271
QY 845 ACGGCACATTCGAAAGCAAGCGTGTCTTCCAGGTGAGCGGGAGCTCTTGGCGGAGCAG 904
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
272 ThrAlaArgArgAlaThrValThrPheGlnThrSerGlyAspSerTyrArgGluGln 291
QY 905 AAGGAGAGCGAGCAGCCATGATGTGGGAATCTGATTTGGTGGTGTGTGTGTGTGTGG 964
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
292 LysGluLysArgAlaAlaMetMetValGlyIleLeuIleGlyValPheValLeuCysTrp 311
QY 965 ATCCCTCTCTCTCCAGGAACTCATCAGCCACCTGTGCTGAGCCTGCCCCCATC 1024
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
312 IleProPhePheLeuThrGluLeuIleSerProLeuCysAlaCysSerLeuProPhe 331
QY 1025 TGGAAAGCATATTTCTGTGGTGTGCTACTCAATTTCTTCAACCCCTGATTTAC 1084
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
332 TrpLysSerIlePheLeuTrpLeuGlyTyrSerAsnSerPheAsnProLeuIleTyr 351
QY 1085 ACAGCTTTTAAAGAACTACACAACTGCTTCAGAGCCTCTTACTAAGCAGAGA 1141
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
352 ThrAlaPheAsnLysAsnTyrAsnAsnAlaPheLysSerLeuPheThrLysGlnArg 370

RESULT 12
US-09-954-342-57
; Sequence 57, Application US/09954342
; Publication No. US20030170838A1
; GENERAL INFORMATION:
; APPLICANT: MISHRA, VISHNU S.
; APPLICANT: SPYTEK, KIMBERLY ANN
; APPLICANT: TAUIER, RAYMOND J.
; APPLICANT: VERNET, CORINE A.
; APPLICANT: COLMAN, STEVEN D.
; APPLICANT: GORMAN, LINDA
; APPLICANT: TCHERNEV, VELIZAR T.
; APPLICANT: MALYANKAR, URIEL M.
; APPLICANT: SHENOY, SURESH
; APPLICANT: PADIGARU, MURALIDHARA
; APPLICANT: GERLACH, VALERIE L.
; APPLICANT: MACDOUGALL, JOHN R.
; APPLICANT: SMITHSON, GLENDA
; APPLICANT: MILLET, ISABELLE
; APPLICANT: PEYMAN, JOHN
; APPLICANT: STONE, DAVID
; APPLICANT: GUNTER, ERIK
; APPLICANT: ELLERMAN, KAREN
; APPLICANT: LI, LI
; APPLICANT: RASTELLI, LUCA
; APPLICANT: ZERHUSEN, BRYAN
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES AND POLYPEPTIDES ENCODED THEREBY
; FILE REFERENCE: 21402-114
; CURRENT APPLICATION NUMBER: US/09/954,342
; CURRENT FILING DATE: 2001-09-17
; PRIOR APPLICATION NUMBER: 60/233,382
; PRIOR FILING DATE: 2000-09-18
; PRIOR APPLICATION NUMBER: 60/240,498
; PRIOR FILING DATE: 2000-10-13
; PRIOR APPLICATION NUMBER: 60/260,284
; PRIOR FILING DATE: 2001-01-08
; PRIOR APPLICATION NUMBER: 60/260,973
; PRIOR FILING DATE: 2001-01-11
; PRIOR APPLICATION NUMBER: 60/264,794
; PRIOR FILING DATE: 2001-01-29
; PRIOR APPLICATION NUMBER: 60/238,398
; PRIOR FILING DATE: 2000-10-06
; PRIOR APPLICATION NUMBER: 60/232,675
; PRIOR FILING DATE: 2000-09-15
; PRIOR APPLICATION NUMBER: 60/274,862
; PRIOR FILING DATE: 2001-03-09
; PRIOR APPLICATION NUMBER: 60/233,801
; PRIOR FILING DATE: 2000-09-19
; PRIOR APPLICATION NUMBER: 60/232,676
```

```
; PRIOR FILING DATE: 2000-09-15
; PRIOR APPLICATION NUMBER: 60/233,960
; PRIOR FILING DATE: 2000-09-20
; PRIOR APPLICATION NUMBER: 60/233,402
; PRIOR FILING DATE: 2000-09-18
; PRIOR APPLICATION NUMBER: 60/233,521
; PRIOR FILING DATE: 2000-09-19
; PRIOR APPLICATION NUMBER: 60/233,522
; PRIOR FILING DATE: 2000-09-19
; PRIOR APPLICATION NUMBER: 60/232,679
; PRIOR FILING DATE: 2000-09-15
; NUMBER OF SEQ ID NOS: 104
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 57
; LENGTH: 370
; TYPE: PRT
; ORGANISM: Mus musculus
US-09-954-342-57

Alignment Scores:
Pred. No.: 4,66e-82 Length: 370
Score: 1452.50 Matches: 298
Percent Similarity: 83.64% Conservative: 19
Best Local Similarity: 78.63% Mismatches: 53
Query Match: 67.43% Indels: 9
DB: 10 Gaps: 3

US-09-976-782-15 (1-1152) x US-09-954-342-57 (1-370)

QY 5 ATGAGAGCGGTAGCTTTCAGTGGCCACCGCGGGTGGCCCTTCCTCGGACCCGAG 64
Db ||||| :|||:||||| ||||| ||||| ||||| ||||| ||||| |||||
1 MetGluValSerAsnLeuSerGlyAlaThrProGlyLeuAlaPheProGlyProGlu 20

QY 65 ACCAGCAGCGGACCCCGGACCCCAAGCCGAGAGGATACTCGGTTTCGACCCCGCGGC 124
Db ||| :||| ||| :|||:||||| ||||| ||||| ||||| ||||| |||||
21 SerCysSer-----AspSerProSerSerGlyArgSerMetGlySerThrProGlyGly 38

QY 125 GCGTCTCTCGCGGCGGAGCGCGCCCTTCTCTGTCTTCCACGCTCCTGTGTGTGACGCTG 184
Db ||||| :|||:||||| ||||| ||||| ||||| ||||| ||||| |||||
39 LeuIleLeuProGlyArgGluProProPheSerAlaPheThrValLeuValThrLeu 58

QY 185 CTAGTGTCTGTATGCTGCCACTTTCCTGTGTGAACCTCTGTGTTCGGTCCACCATCCCG 244
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
59 LeuValLeuLeuIleAlaAlaThrPheLeuTrpAsnLeuValLeuValThrIleLeu 78

QY 245 CGGGTCCGTGCTTCCACCGCGTCCGCGATACTTGTGTGGCTTCGACGCGCTCGGAC 304
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
79 ArgValArgAlaPheHisArgValProHisAsnLeuValAlaSerThrAlaValSerApp 98

QY 305 GAACCTAGTGGCAGCGCTGGCGATGCCACCGAGCTGGCGAGTGAAGTGTGACCGGCGCA 364
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
99 ValLeuValAlaValLeuValMetProLeuSerLeuValSerGluLeuSerAlaGlyArg 118

QY 365 CGTCGGCTCTGGCGCGGAGCTGTGCCACGTGTGATCTCTTCCACCGCGGAGCCTGT 424
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
119 ArgTrpGlnLeuGlyArgSerLeuCysHisValTrpIleSerPheAspVal----- 135

QY 425 CTGTGTGCTCCCGCGGCTCGGGAACGTGGGGCCATCGCCCTGGCGCGGCGGCGCC 484
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
136 LeuCysCysThrAlaSerIleTrpAsnValAlaAlaIleAlaLeuAspArgTyrTrpThr 155

QY 485 ATCACACGGCACCTGCAGCACACGTGCGCACCCGAGCGCGGCTCTGTGTGTCATGATC 544
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
156 IleThrArgHisLeuGlnTyrThrLeuArgThrArgSerArgAlaSerAlaLeuMetIle 175

QY 545 GCGCTCGCGGGTGGCGGCTCATCGCCCTCGCGCGCTGTCTTTGGCCCGGGGC 604
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
176 AlaIleThrTrpAlaLeuSerAlaLeuIleAlaLeuAlaProLeuLeuPheGlyTrpGly 195

QY 605 GAGTGTGCGACGCTCGGCTCCAGCGCTGCCAGCGCTGCCAGTGGAGCGGGAACCTCTATGCGGCC 664
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
196 GluAlaTyrAspAlaArgLeuGlnArgCysGlnValSerGlnGluProSerTyrAlaVal 215
```


Db 312 IleProPhePheLeuThrGluLeuIleSerProLeuCysAlaCysSerLeuProProfile 331
 QY 1025 TGGAAAGCATATTTCTGTGGCTTGGCTACTCCAAATTTCTTCTTCAACCCCTGATTAC 1084
 Db 332 TrpLysSerIlePheLeuTrpLeuGlyTyrSerAsnSerPhePheAsnProLeuIleTyr 351
 QY 1085 ACAGCTTTTAAACAAGTACAAACATGCTTCAAGAGCTCTTTTAAAGCAGAGA 1141
 Db 352 ThrAlaPheAsnLysAsnTyrAsnAsnAlaPheLysSerLeuPheThrLysGlnArg 370

RESULT 14

US-10-109-532A-2
 ; Sequence 2, Application US/10109532A
 ; Publication No. US20030009780A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Allen, Keith D.
 ; TITLE OF INVENTION: TRANSGENIC MICE CONTAINING 5-HT5B
 ; TITLE OF INVENTION: SEROTONIN RECEPTOR GENE DISRUPTIONS
 ; FILE REFERENCE: R-601
 ; CURRENT APPLICATION NUMBER: US/10/109,532A
 ; CURRENT FILING DATE: 2002-03-28
 ; PRIOR APPLICATION NUMBER: US 60/280,553
 ; PRIOR FILING DATE: 2001-03-29
 ; PRIOR APPLICATION NUMBER: US 60/342,472
 ; PRIOR FILING DATE: 2001-12-21
 ; NUMBER OF SEQ ID NOS: 4
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 2
 ; LENGTH: 370
 ; TYPE: PRT
 ; ORGANISM: Mus musculus
 US-10-109-532A-2

Alignment Scores:
 Pred. No.: 370
 Score: 1452.50
 Percent Similarity: 83.64%
 Best Local Similarity: 78.63%
 Query Match: 67.43%
 DB: 14
 Gaps: 3

US-09-976-782-15 (1-1152) x US-10-109-532A-2 (1-370)

QY 5 ATGAGGCGCTAGCTTTCAGTGGCCACCGCGGGTGGCTTCCCTTGGGACCCGAG 64
 Db 1 MetGluValSerAsnLeuSerGlyAlaThrProGlyLeuAlaPheProGlyProGlu 20
 QY 65 ACCAGCAGCGGACCCCGGACCCCAAGCCCGAGAGGATACCTCGGTTTCGACCCCGCGGC 124
 Db 21 SerCysSer-----AspSerProSerSerGlyArgSerMetGlySerThrProGlyGly 38
 QY 125 GCGTCTCGCGGCGCGAGGCGCGCTTCTCTGTCTTCAAGTCTCTGTGGTGAGCGCTG 184
 Db 39 LeuIleLeuProGlyArgGluProProPheSerAlaPheThrValLeuValThrLeu 58
 QY 185 CTAGTCTGCTGATCGCTGCACTTTCCTGTGAACTGCTGTCTCGGFTCAACATCCCG 244
 Db 59 LeuValLeuLeuIleAlaThrPheLeuTrpPheAsnLeuValLeuValThrLeu 78
 QY 245 CCGGTCGCTGCTTCCACCGCGCGCATTAACCTTGTGCTTCAAGTCTCTGTGGCTCGGAC 304
 Db 79 ArgValArgAlaPheHisArgValProHisAsnLeuValAlaSerThrAlaValSerAsp 98
 QY 305 GAACCTAGTGGCGGCTGGCGGATGCCACCGAGCTGGCGAGTGGCTGCGACCGGGCGA 364
 Db 99 ValLeuValAlaValLeuValMetProLeuSerLeuValSerGluLeuSerAlaGlyArg 118
 QY 365 CGTGGCTGCTGGCGGAGCTGTGCGACCTGTGGATCTTCTTCCAGCGCGGAGCGCTGT 424
 Db 119 ArgTrpGlnLeuGlyArgSerLeuCysHisValTrpIleSerPheAspVal----- 135
 QY 425 CTGTGCTGCGCGCGCGCTTGGGAACTGGCGGCGCATCGCTGGCGCGCGCGAGGGGCC 484

Db 136 LeuCysCysThrAlaSerIleTrpAsnValAlaIleAlaLeuAspArgTyrTrpThr 155
 QY 485 ATCACACGGCAGCTCAGCAGCAGCTGGCGCCACCGCGCGCTCGTTGTCTCATGATC 544
 Db 156 IleThrArgHisLeuGlnIleThrLeuArgThrArgSerArgAlaSerAlaLeuMetIle 175
 QY 545 GCGTCTGCGCGCGGTCGCGCTCATCGCGCTTCCGCGCGCTCTCTTTGGCGGGGC 604
 Db 176 AlaIleThrTrpAlaLeuSerAlaLeuIleAlaLeuAlaProLeuLeuPheGlyTrpGly 195
 QY 605 GAGGTGTCGAGCTCGGCTCGGCTCGCAGCTGCGAGGAGCGCGGAGACCTCTATCGCGC 664
 Db 196 GluAlaTyrAspAlaArgLeuGlnArgCysGlnValSerGlnGluProSerTyrAlaVal 215
 QY 665 TTCTCCACCGCGCGCTTCCACCTGCGCTTGGCTGGCTGGCTGGCTGGCTGGCTGG 724
 Db 216 PheSerThrCysGlyAlaPheTyrLeuProLeuAlaValLeuPheValTyrTrpLys 235
 QY 725 ATCTACGAGCGCGCAAGTTTCGTTTCGCGCGCGCGCGAGAGCTGTGCTGCTGGCG 784
 Db 236 IleTyrLysAlaAlaLysPheArgPheGlyArgArgArgAlaValValProLeuPro 255
 QY 785 GCCACATGCGAGTCCCAAGCTAAAGGAGCACCTGATGAGGCTCAAGTGGTGTTC 844
 Db 256 AlaThrThrGln-----AlaLysGluAlaProProGluSerGluMetValPhe 271
 QY 845 ACGGCACATTGCAAGCAACGGTGTCTTCCAGGTGAGCGGGGACTCTCTGGCGGAGCAG 904
 Db 272 ThrAlaArgArgAlaThrValThrPheGlnThrSerGlyAspSerTrpArgGluGln 291
 QY 905 AAGGAGGCGGAGCAGCCATGATGTGGGAATTCGATTGGCTGTGCTGTGCTGTGCTGG 964
 Db 292 LysGluLysArgAlaAlaMetMetValGlyIleLeuIleGlyValPheValLeuCysTrp 311
 QY 965 ATCCCTCTTCTCTGACGGAACCTCATCAGCCCACTGTGCTGAGCTGCGGCTGCCCCATC 1024
 Db 312 IleProPhePheLeuThrGluLeuIleSerProLeuCysAlaCysSerLeuProProfile 331
 QY 1025 TGGAAAGCATATTTCTGTGGCTTGGCTACTCCAAATTTCTTCTTCAACCCCTGATTAC 1084
 Db 332 TrpLysSerIlePheLeuTrpLeuGlyTyrSerAsnSerPhePheAsnProLeuIleTyr 351
 QY 1085 ACAGCTTTTAAACAAGTACAAACATGCTTCAAGAGCTCTTTTAAAGCAGAGA 1141
 Db 352 ThrAlaPheAsnLysAsnTyrAsnAsnAlaPheLysSerLeuPheThrLysGlnArg 370

RESULT 15

US-09-823-187-59
 ; Sequence 59, Application US/09823187
 ; Publication No. US20030096952A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Burgess, Catherine
 ; APPLICANT: Gusev, Vladimir Y
 ; APPLICANT: Liu, Xiaohong
 ; APPLICANT: Majumder, Kumud
 ; APPLICANT: Padigaru, Muralidhar
 ; APPLICANT: Patturajan, Meera
 ; APPLICANT: Shimkets, Richard A
 ; APPLICANT: Spaderna, Steven K
 ; APPLICANT: Spytek, Kimberly
 ; APPLICANT: Taupier, Raymond J
 ; TITLE OF INVENTION: NOVEL PROTEINS AND NUCLEIC ACIDS ENCODING SAME
 ; FILE REFERENCE: 15966-745
 ; CURRENT APPLICATION NUMBER: US/09/823,187
 ; CURRENT FILING DATE: 2001-03-29
 ; PRIOR APPLICATION NUMBER: 60/193,339
 ; PRIOR FILING DATE: 2000-03-30
 ; PRIOR APPLICATION NUMBER: 60/193,205
 ; PRIOR FILING DATE: 2000-03-30
 ; PRIOR APPLICATION NUMBER: 60/195,343
 ; PRIOR FILING DATE: 2000-04-05
 ; PRIOR APPLICATION NUMBER: 60/195,088
 ; PRIOR FILING DATE: 2000-04-06

;; PRIOR APPLICATION NUMBER: 60/195,005
;; PRIOR FILING DATE: 2000-04-06
;; PRIOR APPLICATION NUMBER: 60/195,792
;; PRIOR FILING DATE: 2000-04-10
;; PRIOR APPLICATION NUMBER: 60/196,556
;; PRIOR FILING DATE: 2000-04-11
;; PRIOR APPLICATION NUMBER: 60/197,081
;; PRIOR FILING DATE: 2000-04-13
;; PRIOR APPLICATION NUMBER: 60/197,525
;; PRIOR FILING DATE: 2000-04-14
;; PRIOR APPLICATION NUMBER: 60/197,087
;; PRIOR FILING DATE: 2000-04-14
;; NUMBER OF SEQ ID NOS: 103
;; SOFTWARE: PatentIn Ver. 2.1
;; SEQ ID NO 59
;; LENGTH: 369
;; TYPE: PRT
;; ORGANISM: Rattus norvegicus
US-09-823-187-59

Alignment Scores:
Pred. No.: 1 81e-81 Length: 369
Score: 1443.00 Matches: 236
Percent Similarity: 83.38% Conservative: 20
Best Local Similarity: 78.10% Mismatches: 53
Query Match: 66.99% Indels: 10
DB: 10 Gaps: 4

US-09-976-782-15 (1-1152) x US-09-823-187-59 (1-369)

QY 5 ATGGAGGCGCTGAGCTTTCAAGTGGCCACCGCGCGCTTGCCTTGCCTGGGACCCGAG 64
DB 1 MetGluValSerAsnLeuSerGlyAlaThrProGlyIleAlaPheProProGlyProGlu 20
QY 65 ACCAGCAGGACCGGACCCACCGGACCGGAGGAGTACTCGGTTCGACCCCGAGCGGC 124
DB 21 SerCysSer-----AspSerProSerSerGlyArgSerMetGlySerThrProGly 38
QY 125 GCCGTCTCCGCGCGGCGGCGGCGCTTCTCTGCTTCACTGCTGCTGCTGCTGCTGCTG 184
DB 39 LeuIleLeuSerGlyArgGluProPheSerAlaPheThrValLeuValValThrLeu 58
QY 185 CTAGTGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 244
DB 59 LeuValLeuLeuIleAlaAlaThrPheLeuTrpAsnLeuValLeuValThrIleLeu 78
QY 245 CGGTCCTGCTTCCACCGGCTGCGGATACCTTGTGGTGGCTCGACCGCGCTCTCGGAC 304
DB 79 ArgValArgAlaPheHisArgValProHisAsnLeuValAlaSerThrAlaValSerAsp 98
QY 305 GAACCTAGTGGCAGCGCTGCGATGCCACCGGCGCTGGCGAGTGTGCTGACCGCGCGA 364
DB 99 ValLeuValAlaAlaLeuValMetProLeuSerLeuValSerGluLeuSerAlaGlyArg 118
QY 365 CGTGGCTGTGGGCGGAGCTGTGCCAGTGTGGATCTCTCTTCGACGCGCGGAGCTGT 424
DB 119 ArgTrpGlnLeuGlyArgSerLeuCysHisValTrpIleSerPheAspVal----- 135
QY 425 CTGTGCTGCGCGCGGCTCGGNAACGTGGCGGCTCGCGCTGCGCGCGGCGGCGGCGCC 484
DB 136 LeuCysCysThrAlaSerIleTrpAsnValAlaAlaIleAlaLeuAspArgTrpThr 155
QY 485 ATCACACGGCACCTCAGACACGCTGCGCCACCGCGCGCGCGCTCGTTCATGATC 544
DB 156 IleThrArgHisLeuGlnTrpThrLeuArgThrArgArgAlaSerAlaLeuMetIle 175
QY 545 CGCTCGCGCGGCTGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 604
DB 176 AlaThrTrpAlaLeu---SerAlaLeuIleAlaLeuAlaProLeuLeuPheGlyTrpGly 194
QY 605 GAGGTGTGCGAGCTCGGCTCGAGGCTGCGAGGTGAGCGGGAAACCTCTATGCGCGCC 664
DB 195 GluAlaTrpAlaArgLeuGlnArgCysGlnValSerGlnGluProSerTrpAlaVal 214

QY 665 TTCTCCACCCGCGCGCGCTTCCACCTGCGCTTGGCGTGGTGGTGGTGGTGGTGGTGGTGG 724
DB 215 PheSerThrCysGlyAlaPheTrpValProLeuAlaValValLeuPheValTrpLys 234
QY 725 ATCTACGAGCGCGCAAGTTTCGTTTCGCGCGCGCGCGGAGAGCTGTGCTGCGTTGGCG 784
DB 235 IleTyrLysAlaAlaLysPheArgPheGlyArgArgAlaValValProLeuPro 254
QY 785 GCACCATGCGAGTGGTCCCAAGGTAAAGGAGCACCTGTAGAGCTGAAGTGGTGTTC 844
DB 255 AlaThrThrGln-----AlaLysGluAlaProGlnGluSerGluThrValPhe 270
QY 845 ACGGCACATTGCAAGACGAGTGTCTCCAGGTGAGCGGGGACTCTCTGGCGGAGCAG 904
DB 271 ThrAlaArgCysArgAlaThrValAlaPheGlnThrSerGlyAspSerTrpArgGluGln 290
QY 905 AAGGAGGCGGAGCAGCAGCATGATGCTGGGAATCTGATTGGCGTGTGTGTGTGTGTGTG 964
DB 291 LysGluLysArgAlaAlaMetMetValGlyIleLeuIleGlyValPheValLeuCysTrp 310
QY 965 ATCCCTTCTTCTCCTGACGGAATCATACGCCCACTCTGTGCTGAGCTGCCCGCCATC 1024
DB 311 IleProPhePheLeuThrGluLeuValSerProLeuCysAlaCysSerLeuProIle 330
QY 1025 TGGAAAAGCATATTCTGTGGCTTGGCTACTCCCAATTTCTTCTTCAACCCCTGATTAC 1084
DB 331 TrpLysSerIlePheLeuTrpLeuGlyTyrSerAsnSerPhePheAsnProLeuIleTyr 350
QY 1085 ACAGCTTTTAAAGAACTACAACTGCTTCAAGAGCTCTTTTACTAAGCAGAGA 1141
DB 351 ThrAlaPheAsnLysAsnTyrAsnAsnAlaPheLysSerLeuPheThrLysGlnArg 369

Search completed: August 31, 2004, 20:46:12
Job time : 184 secs

mis Page Blank (uspto)